Charles University

Faculty of Social Sciences

Institute of Economic Studies

MASTER'S THESIS

Good volatility, bad volatility, and the cross-section of stock returns at different investment horizons

Author: **Tony Sako**

Supervisor: **doc. PhDr. Jozef Barunik, Ph.D.**

Academic Year: **2017/2018**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 11, 2018

Signature

Acknowledgments

I wish to express my sincere gratitude especially to doc. PhDr. Jozef Barunik, Ph.D., my thesis supervisor for his unwavering patience, support and guidance during the conception and writing of this thesis report. I also wish to thank my friends and students at Charles University for their support and cooperation in wring this thesis.

Abstract

Starting with the assumption that different investors have different investment time preferences and different risk tolerances within their given investment time-frames, this paper investigates the value of employing multiresolution analysis to model volatility and risk-pricing. In terms of estimation and forecasting performance we were able to reduce by at least half the volatility forecasting errors, with even better results at longer horizons. In regards to risk pricing we learn that extreme aggregate volatility (i.e. tail risk) is priced but regular volatility is not. Additionally we find that whilst aggregate volatility is generally more important over the long-horizon, during periods of market turmoil it is much more significant over the short-horizon. Finally we show that stocks with high sensitivity to aggregate volatility have lower subsequent returns supporting the idea that they become attractive as a hedge against market volatility.

JEL Classification	C12, C13, C21, C22, C31, C32, C51, C52, C53
Keywords	Realized Volatility, Wavelet, Long-Memory Models, Cross-Section, Volatility Forecast, High-Frequency Data

Author's e-mail	tony_sako@yahoo.com
Supervisor's e-mail	barunik@fsv.cuni.cz

Contents

List of Tables

List of Figures

Acronyms

FF Fama-French factors

RV Realized volatility

BPV Bipower variance

SV Semi-variance

PSV Positive semi-variance

NSV Negative semi-variance

APT Arbitrage pricing theory

ACF Autocorrelation function

PACF Partial autocorrelation function

ARCH Autoregressive Conditional Heteroskedasticity

HAR Heterogeneous autoregression

CWT Continuous wavelet transform

DWT Discrete wavelet transform

MO Maximum overlap

HF high frequency

CAPM Capital asset pricing model

ICAPM Intertemporal capital asset pricing model

Master's Thesis Proposal

Author	Tony Sako
Supervisor	doc. PhDr. Jozef Barunik, Ph.D.
Proposed topic	Good volatility, bad volatility, and the cross-section of stock returns at different investment horizons

Motivation It is an established fact that the marketplace consists of many different actors with different perspectives and time-horizons. Additionally, some variables tend to exhibit long memory features which are often not readily evident when analyzing most financial data (e.g. volatility, GDP growth, etc.) using traditional time series methods. Having information that is relevant for the time horizon for which investors are forecasting is important for understanding the specific horizon risk participants face making the correct risk-reward tradeoff decisions.

As already implied, traditional time series methods do not offer an easy way to model the separate behaviors of investors or the different horizon behaviors of various data series. Whilst Fourier methods in the frequency domain are an improvement in allowing for distinguishing between long and short term phenomena and horizons, the localization information is lost due to how the tool is implemented. Wavelet decomposition enables data series to be decomposed into series at different time horizons and offers a way to separate those components of data series that have are long term in nature, and those components that are short term and improves decision-making in the timeframe that investors care about. This allows for example investors when modeling returns to focus on only the risks relevant to their forecast horizon.

The paper will analyze returns and various moments at different horizons to understand at which horizons is contained the most information and evolution over time, and using these data to measure and price risk at different horizons thus allowing an assessment of the risk-reward tradeoff for different specific horizon risk.

Hypotheses

Hypothesis #1: Does wavelet decomposition of volatility measures (volatility, bipower variance, positive semi-variance and negative semi-variance) improve their predictive ability of realized volatility

Hypothesis #2: How do the various moments (volatility, skew and kurtosis) evolve over time at different horizons, and where is the most information created (or contained), if in short or long horizons

Hypothesis #3: Do the moments help to price risk in the CAPM framework in different horizons

Methodology

Data: The data will include 1-minute intraday data over the period from 2005 - 2015 for 29 stocks listed on the Dow Jones and Nasdaq indices from which daily measures (e.g. volatility, skew and kurtosis) will be calculated for the individual stocks as well as equal-weight portfolios to be constructed from the same stocks grouped by sector. In addition daily market prices for the S&P 500 will be used as proxy for the market in the CAPM framework. The risk-free rate as well as the Fama-French factors will be taken from the Ken French library taking daily data. The choice of US data reflects the fact that there is data available going back a long time which allows for analysis of returns data over multiple market cycles which should help to make the results more robust. The data sources will include amongst others Yahoo Finance, the Federal Reserve of St Louis databases and Ken French library and specially obtained files for the intraday data of the 29 stocks used in the study.

Approach: The approach will be to generate daily returns, volatility, bipower variance, semi-variances (positive & negative) as well as skewness and kurtosis from the 1-minute intraday data for each of the stocks. Subsequently these daily measures will be decomposed via discrete wavelet transform into different horizons (e.g. 2-4, 4-8, 8-16 days, etc.). The contribution of each horizon to the risk measure can be evaluated and it's evolution over time assessed potentially grouping into short, medium and long term horizons for simpler analysis.

Modeling: These decomposed horizon data series of the risk measures will also be used in the modeling the pricing of risk in the CAPM framework. For testing volatility forecasting the decomposed data will be used in a HAR framework at different horizons to determine which part(s) of the spectra play an important role in the dynamics of the moments. Similarly, the CAPM model will be extended with the calculated risk measures to determine whether using horizon data improves the ability to price risk. As a robustness check other factors (for example the Fama-French 3-factors) will be included in the model to see whether the coefficients remain economically and statistically significant.

Expected Contribution Heterogeneous autoregression (HAR) models are already well-known in finance literature. This paper will test the hypothesis that there is a benefit over current methods in applying the HAR model to different horizons. Similarly, the paper will test whether the CAPM framework could be enhanced by incorporating the moments of returns at different horizons in the framework in pricing risk. If successful, the model will better explain and forecast moments in comparison to HAR and for investors such an approach would provide an additional tool in the measurement and pricing of risk relevant to their investment horizon.

Outline

1. Motivation: there are various ways of measuring and pricing risk, however they do not take into account the different horizons of investors or the long memory inherent in some types of asset prices. Using wavelet transform overcomes these issues and by incorporating various risk measures in the HAR model and CAPM framework it is possible to improve both the models for measuring and pricing of risk.
2. Studies on measuring and pricing risk: I will briefly describe the existing literature on measuring and pricing risk in the time and time-scale domains.
3. Data: will explain how I will construct the data (in case of sector portfolios) and calculate the various risk measures and the theoretical underpinning for the calculated risk measures.
4. Methods: I will explain the concepts of wavelets, HAR model and CAPM framework, weighted least squares and pooled panel which are expected to feature in the analysis
5. Results: I will discuss my baseline regressions and robustness checks.
6. Concluding remarks: I will summarize my findings and their implications for future research.

Core bibliography

Andrew J. Patton and Kevin Sheppard, 2015. Good volatility, bad volatility: signed jumps and the persistence of volatility. The Review of Economics and Statistics 2015 97:3, 683-697.

Martijn Boons and Andrea Tamoni, 2016. Horizon-specic macroeconomic risks and the cross section of expected returns.

Andrea Tamoni, 2011. The Multi-Horizon Dynamics of Risk and Returns. Unpublished manuscript.

Marco Gallegati, 2008. Wavelet analysis of stock returns and aggregate economic activity. Computational Statistics & Data Analysis 52 (2008) 3061 - 3074.

Diego Amaya, Peter Christoffersen, Kris Jacobs and Aurelio Vasquez, 2015. Does Realized Skewness Predict the Cross-Section of Equity Returns. Journal of Financial Economics. Volume 118, October 2015, Pages 135-167.

Author

Supervisor

Chapter 1

Introduction

Uncertainty plays a key role in finance in terms of shaping the landscape of investment opportunities. This uncertainty is typically quantified in terms of second (variances and covariances) or higher moments; therefore a good understanding of both these moments is important in identifying and managing risk. Volatility for instance is a central topic in finance from asset pricing, derivatives valuation, asset allocation, portfolio optimization, and risk management. Equally importantly, covariance is a key factor in determining the systemic risk of an asset and what impact the asset would have in a portfolio. Added to this is the fact that globally there are a large and growing number and volume of traded products specifically geared towards volatility which underscores the importance of being able to measure and forecast uncertainty with high precision.

In recent times high frequency (HF) intraday returns have been used with great success to estimate volatility with great precision. A key advantage of this approach is that unlike other approaches diffusion volatility does not require a long series of data in order to estimate very precise estimates. If the data is sampled sufficiently frequently it can be possible to obtain very good estimates even with a short series of data. Taking advantage of this and other attractive features of HF data, we use 1-minute intraday data in constructing various estimates of the underlying volatility.

However, whilst realized volatility measures can provide a very precise measure of volatility ex-post and is very useful for many situations where historical volatility is a key input (e.g. for VaR analysis), it is not sufficient for all investor needs. For many investors who are generally forward-looking an important aspect of risk management extends beyond simply extrapolating historical volatility developments into the future, but also being able to more accurately model and forecast future volatility that they will face. This paper examines therefore two main aspects in relation to volatility that have not received much attention. Investigating whether stock volatility estimates and forecasts could be improved by using some of the realized volatility estimators within a HAR framework and determining, within the HAR framework, what the best estimator for stock market volatility is. Using the HAR analysis has the particular advantage that it enables modeling of the long-term memory feature of volatility without requiring at the same time a long history of data to work with, so is ideally suited especially for high-frequency intraday data.

We suppose, as is generally accepted view, that markets consist of many heterogeneous agents each with their own risk preferences and investment horizons from short term day traders to institutional investors. An example to demonstrate this, if we consider that institutional investors have a long-term horizon and base much of their decision-making on trends and developments in macroeconomic factors, then we might expect for there to be a stronger correlation between long-term market performance and macroeconomic developments which are long-term by nature than there would be between short-term market performance and long-term economic developments. In such a case an analysis that was able to extract the signal dynamics at different horizons would be very useful in making sense of such relationships that are strong over certain horizons than over other horizons. And from the perspective of the long-term investor, short-term volatility may be a mere distraction as their focus is on the underlying macroeconomic trends which are generally slow-moving and long-term in duration. A volatility measure that is stripped of the short-term perturbations might present a truer picture of the underlying long-term uncertainty for such an investor.

Traditional time series methods do not offer an easy way to model the separate the different dynamics at different horizons and it is for this purpose that we will deploy wavelet transforms in this paper. This paper will focus on using wavelets to construct volatility measures at different horizons and investigating whether the modeling and forecasting performance would be enhanced. Whilst realized volatilities and HAR, or realized and multiresolution framework have been studied, what this paper does differently is analyze the HAR model within a multi-horizon framework. We construct a horizon-specific measure of volatility and compare how well our estimate does compared to the observed volatility as well as comparing various estimators with the aim to improve the forecasting precision.

In addition to the improvement in measurement and forecasting of individual asset volatility, another question that this paper investigates is whether aggregate volatility is a risk factor and whether or not it is priced in the cross-section of returns. The question whether aggregate volatility is priced in the cross-section of returns is rather unclear and a number of studies have published conflicting results either in terms of statistical and economic significance, or in terms of the sign of the coefficient. What we do differently in this paper is to carry out the analysis in a horizon-specific manner to understand whether in a horizon-specific framework aggregate volatility is a risk factor and whether investors with different horizons face the same aggregate volatility risk. And lastly, observing that market volatility is positively correlated with future uncertainty about returns, which is also theoretical founded, there is the question whether assets with a high correlation to market volatility might behave as a hedge against future uncertainty. To understand this further the paper will investigate whether assets with a high correlation to market volatility have lower average future returns and high contemporaneous average returns.

The objective of this thesis therefore is to examine these questions which have already been posed in this introductory text in relation to improving the precision in measurement of volatility and providing a horizon-specific measure of volatility using realized volatility methods, HAR analysis and wavelet analysis. Additionally analyzing aggregate volatility as a systemic risk factor candidate and whether the risk premia associated with aggregate volatility are the same for investors with different horizons.

The rest of the work is structured as follows: Chapter 2 describes an overview of the relevant literature on the topics of this paper namely volatility, cross-section of expected returns, wavelet transforms applied to financial data. Chapter 3 provides the theoretical overview covering the main topics covered in the paper. Chapter 4 gives an overview of the data description including the summary statistics, correlations and where data were transformed or new variables created descriptions of these transformations are provided. Also, provides some additional details around methodology. Chapter 5 presents the empirical results of the volatility analysis. Chapter 6 presents the empirical results from the cross-section analysis. Chapter 7 summarizes the findings and recommendations for future research.

Chapter 2

Literature Review

2.1 Volatility

There is a large volume of literature modeling the persistence of financial markets volatility using ARCH/GARCH and stochastic volatility models. Most of the early studies had documented good in-sample fits but quite poor out-of-sample forecasts. Andersen & Bollerslev (1998) showed however that with good model specification volatility forecasts can be improved. Their main result however hinged on the use of frequently sampled data pointing to the effective use of high-frequency intraday data in modeling of volatility.

Andersen et al. (2000) introduced a new volatility measure named realized volatility and theoretical framework for integration of high frequency data in the measurement and forecasting of volatility. Two very useful properties of these volatility measures introduced include the fact that by construction they are model-free thus alleviating the issue of model specification. Secondly, they make it possible to estimate volatility for almost any horizon thus avoiding the need for long history data in order to pick out the trend.

Applying the newly defined realized volatility measures to highly liquid currency (DM/$ and $/Yen exchange rates) intraday data with 10 years of 5-minute, the study finds that whilst daily variance, standard deviation and correlations are right-skewed and leptokurtic, typical of financial data series, the distributions of logarithmic standard deviations are approximately normally distributed. And in a departure from earlier work show they that volatility persistence does not decrease quickly with increasing horizon and that the clustering effects are still strongly present even on weekly or monthly data.

In subsequent work Andersen *et al.* (2001) extend the analysis to individual stocks listed on the Dow Jones Industrial Average index evaluating both conditional and unconditional volatility distributions. The study finds that for the unconditional distributions, similar to the case for currency exchange rates, the realized variances and standard deviations are non-normal and right skewed but the logarithmic counterparts are approximately normally distributed. Returns distributions are found to be leptokurtic, but the returns normalized by standard deviation are approximately normal. Similarly, whilst the covariances are right-skewed, the realized correlations are approximately normally distributed.

In terms of conditional distributions they find that the realized variances fluctuate significantly over time and are best described by a mean-reverting long memory process with d parameter of approximately 0.35. Interestingly they find that whilst the much discussed leverage effect (i.e. asymmetric relation between past sign of returns and future volatility) is statistically significant, economically it's relative unimportant, at least at the individual stock level. They conclude that due significant asymmetries at the market level reported in other papers, the best explanation cannot be a leverage effect but instead should be due to a volatility feedback effect.

Further extending the work on realized variance measures, Barndorff-Nielsen & Shephard (2004b) introduced the realized bipower variation which they showed estimates integrated variance in stochastic volatility models, thus providing a model-free alternative to realized variance. The difference between realized variance and realized bipower variation estimates the quadratic variation of the jump component. Therefore, making it possible to separate quadratic variance into the continuous and jump components.

Barndorff-Nielsen *et al.* (2009) note that prior to their work economists had been interested in measuring downside risk and had until that time come up with a number of measures including semivariance, value at risk and expected shortfall. Building on the foundations already laid on realized variance measures, they introduced a new measure of downside and upside risk called realized semivariance. They showed that the new measure was outperforming the GARCH and GJR models and could also be incorporated in these models thereby enriching them.

In the study by Patton & Sheppard (2015), they showed that future volatility is more strongly related to volatility of past negative returns than that of past positive returns. Their investigation also showed that the impact of a price jump on future volatility is dependent on the sign of the jump with negative jumps associated with higher future volatility than positive jumps. This result was in contrast to earlier work by Andersen *et al.* (2007) which found that jumps had limited predictive value on future volatility. The difference being that earlier work focused on unsigned jumps but by separating into positive and negative jumps their predictive value can be unlocked. When it comes to forecast performance they showed that the model has better out-of-sample forecast performance for forecast horizons between 1 to 3 months.

In a study on a large cross-section of stocks (nearly twenty thousand stocks) covering a period of twenty years, Bollerslev *et al.* (2016) decompose realized volatility into positive and negative semi-variance (i.e. good and bad volatilities). In the cross-sectional regressions they sort the stocks into portfolios based on each stock's good minus bad volatility. Their study finds that the differences in returns remain statistically and economically significant even after controlling for known explanatory variables. They show that stocks with high good-to-bad volatility ratio earn higher weekly returns than stocks with low good-to-bad volatility ratio (i.e. realized signed jumps). They also relate the realized signed jumps to other firm characteristics such as size, liquidity, etc. The t-statistic of 9.66 they obtained also exceeds the more stringent hurdle rates for judging statistical significance in cross-sectional studies that was recently advocated by Harvey (2016).

2.2 Wavelets In Finance

According to Tamoni (2011) regular time series that are exclusively focused on a scale do not have the ability to explain the underlying data generating process. The issue is one of aggregation where it's not feasible to incorporate long-term effects into a model that is focused on short-term effects. It is therefore a challenge to find models that are able to incorporate multi-horizon data.

Wavelet analysis is a more recent approach to analyzing a time series and it can be seen as an extension of Fourier analysis. Both techniques allow a signal to be transformed from the time domain into frequency domain where some features may be more easily highlighted or the easier to extract and work with.

Addison (2002) notes that Fourier analysis has some disadvantages as compared to wavelet analysis. Firstly, Fourier analysis requires that the data series is stationary. In terms of financial data series this is a big challenge. Secondly, due to the fact that the basis functions have infinite support, time localization data is lost in the Fourier transform. However with wavelet analysis localization data in both time and frequency is preserved. Wavelet analysis also does not require the data series to be stationary so it's quite ideal for financial data.

The field of wavelet analysis really started to take off after the work of Mallat (1989) who was able to unite a number of different topics from signal processing and image processing to come up with the concept of a multiresolution analysis. The pyramid algorithm he introduced for implementing the discrete wavelet transform was very efficient and also succeeded in making the wavelet transform more practical.

The term wavelets stands for small waves as they are compact and oscillatory. They are a type of function which is used to decompose a signal into different components corresponding to different time-scales. They are a type of function created from a prototype called the "mother" wavelet that can be scaled (dilated or compressed) and translated in order to zoom in on different levels of detail. Building on the work of Mallat, Daubechies (1992) introduced a new family of wavelets which are quite useful in analysis of time series as they were smoother and more efficient and also allow a very accurate alignment of the transform coefficients and the original time series.

Ramsey (2002) provides a listing of some of the applications in economics and finance and demonstrates that wavelets have already been quite useful in proving valuable insights. A time series may contain features which are prominent at some time-scales but not at others and using multiresolution analysis it's possible to separate out those features Tamoni (2011). Chaudhuri & Lo (2016) argue that economic shocks can have diverse effects on financial market dynamics at different time horizons and show that decomposing asset-return variances, correlations, alphas, and betas into distinct frequency components can identify the relative importance of specific time horizons in determining each of these quantities as well as in constructing mean-variance-frequency optimal portfolios.

Gencay *et al.* (2005) proposed a new approach to estimating systematic risk (i.e. market beta) using wavelets to decompose a given time series on a scale-by-scale basis. They found that beta measure became more important as the scale increased. This has at least two implications, firstly that systemic risk of an asset as measured by beta is different at different horizons. And secondly it suggests that beta might be more relevant in predicting risk on medium- to long-run horizons as compared to short time horizons. In this paper we investigate whether a similar dynamic also exists for aggregate volatility where the volatility beta is different at different horizons and if it becomes increasingly important as the scale increases.

Gencay *et al.* (2001) apply wavelet analysis on realized volatility using currency (DM/$ and $/Yen exchange rates) intraday data with 10 years of 5-minute. They find there is a smaller degree of persistence in intraday volatility as compared to volatility at one day and higher scales and correlation is also lowest intraday as opposed to longer timeframes. As we also conduct a wavelet analysis of realized volatility on 1-minute data in this paper this finding does have an interesting implication for this paper in that since we're using HAR analysis to model the long-memory features of volatility then we should expect to have a much better model fit at higher scales than at the lower scales.

Extending the use of wavelet analysis in decomposing HF variance estimators, Barunik & Vacha (2015) estimate integrated variance in the presence of noise using the MODWT transform, which is more efficient, and also employing the Daubechies wavelets. They're able to estimate very precise estimates that were also robust to noise and jump level, data generating process or investment horizon. Using the result as inspiration we also employ the MODWT wavelet transform as well as utilizing the Daubechies family of wavelets in the transform in the search for more precise estimates of realized volatility.

2.3 Volatility and Cross-Sections of Returns

Arbitrage pricing theory (APT) was introduced by Ross (1976) in part as a response to the issues that were being raised surrounding the Capital asset pricing model (CAPM) which was the predominant theory for asset pricing. CAPM had proposed that the only relevant systemic risk was the market returns but many studies had found that it was not adequately explaining the cross section of returns. With the APT Ross argued that there is a common component in the movement of stock returns and this tendency for asset returns to move together could be modeled by statistical factor decomposition. Furthermore, if a stock's returns can be synthesized by a portfolio of factors, then by the law of one price the price for the stock returns should be derivable from the factor return prices.

Ross's model requires an exact factor structure (i.e. idiosyncratic errors in the time series regression for the factor betas are zero), however Cochrane (2001) notes that actual returns do not display an exact factor structure. Chamberlain & Rothschild (1983) extend Ross's work and show that an approximate factor structure (when errors are small) is also sufficient for Ross's result. Further to that Cochrane notes that the APT regressions of the form Ross suggested typically have high R^2 (i.e. small errors). Thus even if an individual stock may have a high idiosyncratic errors, within a well-diversified portfolio the idiosyncratic errors should make only a small contribution overall.

Fama & French (1992) introduced their FF 3-Factor model, generally categorized as an APT model, which has been quite successful in explaining the cross-section of returns. Due however to a number of observed anomalies, there has been an intensive search in academia and industry for additional factors that could help explain away those anomalies leading to an explosion in the number of discovered factors. Harvey (2016) have counted more than 300 papers written on the subject and make the point that given the number of factors that have been discovered, more stricter criteria on the p-values are required on any new factors. They do acknowledge however that not all factors need be treated the same and that factors derived from theory ought to have a lower hurdle than factors derived purely from empirical exercise.

In the section on cross-section of returns we investigate whether aggregate volatility is a factor and whether it is priced the same at all horizons. Thus this paper contributes to that search for additional factors to the known factors. We consider that the case for aggregate volatility as a factor has strong theoretical foundations and there does not need to be additional stringent statistical criteria applied. The work of Merton (1973) and Ross (1976) underline that risk premia are associated with the conditional covariances of asset returns and factors which are linked with the performance in returns.

A number of studies within the ICAPM literature highlight the importance of volatility in forecasting future returns and uncertainty. Campbell (1996) finds that investors care about not just market returns but also the changes in forecasts of future returns. Chen (2002) extends Campbell's model to find that an asset's expected returns depend also on changes in forecasts of future market volatility. Ang *et al.* (2006) bases his empirical study on this theoretical foundation and finds that stocks that have a high sensitivity to aggregate volatility have low average returns and that stocks with high idiosyncratic volatility (within a FF 3-factor framework) have even lower average returns. We conclude from the literature that there is merit to analyzing further aggregate volatility as a factor and the theoretical foundation is also strong.

Chapter 3

3 Theoretical Overview

3.1 Volatility

In this section we present formally the volatility estimators used in the paper. Generally stock prices are considered to follow a jump-diffusion model and also containing noise as

$$y_t = p_t + \epsilon_t$$
$$\text{where } \epsilon_t \approx N(0, \sigma^2)$$

and

$$dp_t = \mu_t dt + \sigma_t dW_t + c_t dJ_t, \tag{3.1}$$

where μ_t denote the drift term, σ the diffusive volatility process, W is a standard Brownian motion, and J is a pure jump process. In the jump process c_t refers to the size of the jump and dJ_t is a count where that $J = 1$ if there is a jump at time t (and 0 otherwise).

The logarithmic jump diffusion price, where the unit time-interval corresponds to a trading day, is given by

$$\int p_T = \int_0^T \mu_T dt + \int_0^T \sigma_T dW_T + J_T \tag{3.2}$$

Assuming that high-frequency intraday prices $p_t, p_{t+1/n}, \ldots, p_{t+1}$ are observed at $n + 1$ equally spaced times over the trading day $[t, t + 1]$. Then the corresponding logarithmic discrete-time return over the ith time-interval on day $t+1$ is given by

$$r_{t+i/n} = p_{t+i/n} - p_{t+(i-1)/n} \tag{3.3}$$

The daily realized variance, an estimate of the quadratic variation, is then defined by the summation of the intra-day high-frequency squared returns as noted in Barndorff-Nielsen & Shephard (2004a) and Barndorff-Nielsen & Shephard (2004b). The quadratic variation of the process therefore is given by

$$RV_t = \sum_{i=1}^{n} r^2_{t-1+i/n} \tag{3.4}$$

The integrated variance of the process is estimated using realized Bipower variance (BPV) defined by

$$BPV_t = \frac{\pi}{2}\frac{n}{n-1}\sum_{i=2}^{n} |r_{t+i}||r_{t+(i-1)}| \tag{3.5}$$

Andersen *et al.* (2003) show that the realized variance converges to the quadratic variation; and is comprised of separate components due to the continuous and jump price increments, as n approaches infinity

$$\lim_{n\to\infty} RV_t \to \int_{t-1}^{t} \sigma_s^2 ds + \sum_{t-1\leq\tau\leq t} J_\tau^2 \tag{3.6}$$

Therefore, using the above results and given appropriate regularity conditions (Clements & Liao 2013), we have

$$\lim_{n\to\infty} BPV_t \to \int_{t-1}^{t} \sigma_s^2 ds \tag{3.7}$$

Implied in the above result is that subsequently the difference between the two estimators (RV and BPV) can be used to estimate the contribution of the jump component to volatility. However, there is a caveat. Due to the fact that as the sampling period gets ever smaller the size of individual price increments in each sampling period is also getting smaller such that eventually measurement error becomes a significant component of the observed price increment. Therefore in praxis there is a constraint as to how small the sampling period can be. Furthermore, small jumps should be treated as measurement error and only sufficiently large jumps should be identified as jump increments.

In order to test whether a sufficiently large jump has occurred a statistical test on the jump component has been introduced such that,

$$J_t^2 \equiv I_{Z_t > \Phi(\alpha)}(RV_t - BPV_t), \tag{3.8}$$

where $\Phi(\alpha)$ is a critical value from the $N(0,1)$ distribution and Z_t is a statistic used for testing for the presence of jumps under the null hypothesis of no jumps (Hausman test). The Z_t statistic is defined as

$$Z_t = \frac{(RV_t - BPV_t)/RV_t}{\sqrt{\frac{(\pi/2)^2+\pi-5}{n} * max(1, \frac{TQ_t}{BPV^2})}} \tag{3.9}$$

where TQ_t (realized tripower quarticity measure) is the fourth moment and $max(1, \frac{TQ_t}{BPV^2})$ is its small sample refinement.

$$TQ_t = \frac{\Delta^2/(\Delta-4)}{0.8313^3} \sum_{i=5}^{n} |r_{t-1+(i-4)\Delta}|^{4/3} |r_{t-1+(i-3)\Delta}|^{4/3} |r_{t-1+(i-2)\Delta}|^{4/3} \tag{3.10}$$

where $\Delta = 1/n$ is the fraction of the trading session associated with the sampling frequency. If the null hypothesis of no jumps is not rejected then the "jump" component is considered as part of the continuous increment so that the continuous and jump components always add up to the realized volatility.

It's important to note that the realized volatility derived above does not distinguish between upside volatility and downside volatility (i.e. good volatility and bad volatility). To model this phenomena (Barndorff-Nielsen *et al.* 2009) proposed the realized up and down semi-variance measures to decompose realized variance into positive semi-variance and negative semi-variance which are associated with positive and negative returns respectively.

$$RV_t^+ = \sum_{i=1}^{n} r_{t-1+i/n}^2 1_{\{r_{t-1+i/n}>0\}} \tag{3.11}$$

$$RV_t^- = \sum_{i=1}^{n} r_{t-1+i/n}^2 1_{\{r_{t-1+i/n}<0\}} \tag{3.12}$$

where the semi-variances add up to the realized variance, i.e. $RV_t = RV_t^+ + RV_t^-$.

Furthermore, (Bollerslev *et al.* 2016), find that in the limit the semi-variances converge to

$$RV_t^+ \rightarrow \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds + \sum_{t-1\leq\tau\leq t} J_\tau^2 1_{(J_\tau>0)} \tag{3.13}$$

$$RV_t^- \rightarrow \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds + \sum_{t-1\leq\tau\leq t} J_\tau^2 1_{(J_\tau<0)} \tag{3.14}$$

which basically states that the continuous part of both the positive and negative semi-variances converge to one half of the integrated variance and the only difference between the two is in the jump component.

As a final component in estimating realized volatility is what Patton & Sheppard (2015) define signed jumps as

$$SJ_t^2 \equiv RV_t^+ - RV_t^- \rightarrow \sum_{t-1\leq\tau\leq t} J_\tau^2 1_{(J_\tau>0)} - \sum_{t-1\leq\tau\leq t} J_\tau^2 1_{(J_\tau<0)} \tag{3.15}$$

In the regressions in this paper we will use both the unsigned jump J^2 component as well as the signed jump SJ^2 component together with BPV in estimating realized volatility. As a final note, the signed component component as defined here may not be used in conjunction with SV in a regression due to the problem of perfect collinearity. Therefore, for realized volatility regression on SV we include only the J^2 component if necessary.

3.2 Wavelet Transforms

This section provides an overview of the theoretical foundations of wavelet analysis. Broadly speaking wavelet analysis is a tool that allows for a signal to be decomposed into a series of signals of different levels of detail. This allows for the possibility of zooming in and studying the details or zooming out and observing the broader trends in the signal.

Wavelet analysis is based on the well-known Fourier analysis which has found use in many fields of science and engineering. However a wavelet transformation can be localized in both time and frequency whereas Fourier transform loses the time localization. In a Fourier transformation a given signal can be represented as a sum of sine and cosine functions at different frequencies.

Wavelet transformation performs a similar decomposition but a key difference is that wavelet transform uses wavelets instead of sines and cosines. Wavelets can be understood as "small" waves which have compact support and therefore finite energy and duration as opposed to sines and cosines which have infinite energy and have support between plus and minus infinity.

The result is that whereas Fourier transformation maps a one-dimensional function of a continuous variable into a one-dimensional sequence of coefficients at different frequencies, the wavelet transformation on the other hand maps a signal into a two-dimensional array of coefficients. It is primarily due to this two-dimensional transformation which makes the wavelet transformation capable to localize the signal transform in both time and frequency. By being able to compress or stretch wavelet support, the wavelet transform is able respectively to zoom in and capture short-run phenomena (e.g. spikes, jumps and transitions) or zoom out and capture trends and other long-run phenomena.

Wavelet transform decomposes a time series using elementary functions called mother wavelets that are expressed as a function of translation τ (time) and scale s (dilation), which is related to frequency. Mathematically, these wavelets result from a mother wavelet are defined as

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}}\psi(\frac{t-\tau}{s}) \tag{3.16}$$

In order to be considered a mother wavelet the function $\psi(t)$ must fulfill several conditions. These are described in detail in (Percival & Walden 2000) but the key criteria are admissibility and regularity.

Admissibility condition states that a function $\psi(t)$ satisfying the condition

$$\int \frac{|\Psi(w)|^2}{w}\, dw < \infty, \tag{3.17}$$

where $\Psi(w)$ is the Fourier transform of the $\psi(t)$, can be used to both transform and then reconstruct a signal without loss of information. This condition implies that

$$\left.|\Psi(w)|^2\right|_{w=0} = 0$$

so that the function $\psi(t)$ should have a band-pass spectrum. Additionally, this also implies that the average value of the function is zero so that

$$\int \psi(t)\, dt = 0$$

Both of these implications above result in a mother wavelet function that is both a wave and has finite support.

The regularity condition concerns with the constraints imposed on the wavelet function to have some smoothness and concentration in both time and frequency domains in order to ensure that the wavelet coefficients will decay quickly as the scale increases. The decay is dependent on the number of vanishing moments whereby if a wavelet has n vanishing moments, then the wavelet transform coefficients $W_x(\tau, s)$ will decay as fast as s^{n+2} for a smooth signal $x(t)$.

Wavelet transform come in two basic varieties, namely continuous and discrete. The Continuous wavelet transform (CWT) of a process, x(t), with respect to a given wavelet function $\psi(t)$, is given by the convolution of the process and wavelet function as

$$W_x(\tau, s) = \int_{-\infty}^{+\infty} x(t)\psi^*(t)\,\mathrm{dt} = \int_{-\infty}^{+\infty} x(t)\frac{1}{\sqrt{s}}\psi^*(\frac{t-\tau}{s})\,\mathrm{dt}, \tag{3.18}$$

where $\psi^*(t)$ is the complex conjugate

As we use the discrete wavelet transform in this paper we do not discuss further the CWT and instead focus on the discrete transform.

3.2.1 Discrete Wavelet Transform

The CWT as described has a lot of redundant information and is not nearly as practical due to the fact there are an infinite number of wavelets to transform and also due to the fact that the wavelet transforms have no analytical solution and therefore have to be solved numerically. Discrete wavelet transform (DWT) have been introduced to address this problem. To avoid data redundancy DWT are scaled and translated in discrete steps instead of continuously. This is achieved by a slight modification of the wavelet function

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}}\psi(\frac{t-2^jk}{2^j}) \tag{3.19}$$

corresponding to a dyadic sampling. To avoid that an infinite number of scales is required in order to cover all frequencies up to zero, the scaling function is used in combination with the wavelet functions. The scaling function is a low-pass filter which can cover all the frequencies not covered by the wavelets. Therefore combining the low-pass scaling function with the wavelet functions which are band-pass filters a filter bank can be constructed which allows to capture all the frequencies.

Percival & Mofjeld (1997) show that because the DWT matrix W is orthonormal by construction, this implies that

$$||\mathbf{W}||^2 = ||\mathbf{X}||^2 \equiv \sum_{t=0}^{N-1} X_t^2$$

Decomposing $\mathbf{W}$ into the wavelet $\mathbf{W_j}$ and scale $\mathbf{V_J}$ subvectors gives

$$||\mathbf{X}||^2 = ||\mathbf{W_j}||^2 + ||\mathbf{V_J}||^2, \tag{3.20}$$

where $||\mathbf{W_j}||^2$ represents the contribution to the squared norm of $\mathbf{X}$ due to changes at scale j and $||\mathbf{V_J}||^2$ represents the contribution due to variations from the rest of the scales (i.e. scales $J+1$ and higher). This result implies that the signal variance is exactly decomposed into the wavelet variance at each scale (Serroukh *et al.* 2000).

Lindsay *et al.* (1996) also show that the orthonormal wavelet decomposition of a data series as is the case with the DWT leads to a natural partition of the variance by scale. The variance may be expressed in terms of the wavelet coefficients

$$\hat{\sigma_x}^2 = \frac{1}{N} \sum_{j=1}^{J} \sum_{k=1}^{n_j} D_{j,k} \tag{3.21}$$

And the contribution to the total variance from the variance $\hat{\sigma_{x,j}}$ of each scale j is

$$\hat{\sigma}^2_{x,j} = \frac{n_j}{N} \left(\frac{1}{n_j} \sum_{k=1}^{n_j} D_{j,k} \right) = \frac{1}{2^j} \hat{\sigma}^2_{D,j} \tag{3.22}$$

$$\hat{\sigma}^2_{x,j} = \frac{1}{2^j} \hat{\sigma}^2_{D,j} \tag{3.23}$$

where $n_j = N/2^j$ and $\hat{\sigma}^2_{x,j}$ is the sample variance of the wavelet coefficients

D_j, at scale j. The Maximum overlap (MO) sample estimate of the wavelet variance at scale j is given by

$$\tilde{\sigma}^2_{x,j} = \frac{1}{2^j \tilde{n_j}} \sum_{k=1}^{\tilde{n}_j} d_{j,k}, \tag{3.24}$$

where $\tilde{n_j}$ is the number of MO coefficients at scale j.

3.3 Volatility and Cross-Sections

In this section we will motivate for the use of the APT and argue for aggregate volatility as a common factor in the APT model. The key insight from the CAPM model was that risk arises from the covariance of an asset with the market. In the CAPM model systemic risk was the market factor and investors ought to be to be compensated proportional to how the asset covaries with the market since this risk cannot be diversified away. Due however to the failure in explaining the cross-section of returns, multi-factor models have been developed which incorporate other state variables beyond just the market factor. The multi-factor models developed by Ross (1976) and Merton (1973) expand on the insights of the CAPM beyond the market factor as the sole systematic risk showing that risk premia are associated with the covariance of an asset with the variations in state variables that describe the time variation of investment opportunities (Ang *et al.* 2006).

APT models start from a statistical perspective by arguing that stock returns tend to move together and therefore there is a common component affecting stock returns. Therefore the APT defines a mathematical model that can adequately describe this tendency of stocks to move together via a statistical factor decomposition. The concept being that it should be possible to replicate the returns of a stock by constructing a suitable portfolio of factors. The APT however, unlike the CAPM, does not specify what the factors should be.

Cochrane (2001) does however offer a clue how to evaluate factors. As the intuition behind the APT is that stock returns can be synthesized by a factor portfolio, purely idiosyncratic risk should not be associated with any risk premium and the expected returns on a stock should be related only to the covariance of the stock returns with the common factors. So whilst this approach does not state what the factors should be upfront, we have a way to validate after the fact we have selected good factors by checking whether idiosyncratic errors are priced. By contrast the factor returns should be priced in the cross-section of expected returns.

Responding to the failures of the CAPM to explain a number of anomalies, Fama & French (1992) introduced the now famous three-factor model with size, value and market return as factors which has been quite successful in explaining the cross-section of returns. And in line with many papers written in this field we will also use the Fama-French model as the basis for our cross-section analysis. However many so-called "anomalies" associated with the Fama-French model have continued to surface and the search for additional factors continues. Fama and French have also since augmented their model to include an additional two factors Fama & French (2016). And many others have also found other factors that help explain the cross-section of returns. In a similar manner we investigate the use of aggregate volatility as a new factor.

The theoretical inspiration for aggregate volatility as a factor arises from the Intertemporal capital asset pricing model (ICAPM) literature. Campbell (1996) approaching from the perspective of investor utility, shows that investors care about risks both from the market return and from changes in forecasts of future market returns. Therefore assets that covary positively with expected future market returns have a higher premium as they as they are viewed as risky. This is directly tied to the question whether volatility is a relevant risk factor since volatility is positively associates with future expected returns.

Chen (2002) extends Campbell's model to allow for time-varying covariances and stochastic market volatility and shows that an asset's expected return depends on risk from the market return, changes in forecasts of future market returns, and changes in forecasts of future market volatilities. This model implies that if an asset's returns covary positively with an asset that positively forecasts future market volatilities, then that asset's expected returns will be lower. The reasoning being that an investor will be unhappy with news that future returns will be lower as their consumption will be lower and will therefore prefer stocks that do well in those situations to hedge their reinvestment risk. Thus by demanding more of such assets they raise the asset prices thereby lowering future expected returns.

Ang *et al.* (2006) investigated this phenomenon and found that stocks with high sensitivities to innovations in aggregate volatility have low average returns. Their approach was sorting stocks based into five portfolios based on the aggregate volatility betas from the time series regression and then comparing average returns between the highest and lowest sensitivity portfolios. In their time series regression however they only had aggregate volatility and market return as factors. What this paper does differently is include in the time series regression all the Fama-French to control for all the important known factors.

Turning now to a formal description of the APT model we use in this paper. The APT states

$$r_i = a_i + \sum_{k=1}^{N} \beta_{i,k} \tilde{f}_k + \epsilon_i \tag{3.25}$$

where r_i are the excess returns for asset i, N is the number of factors, $\beta_{i,k}$ is the beta for ith on factor k, $\tilde{f}_k$ is the kth factor with $\tilde{f} \equiv f - E(f)$. The factors are typically defined as innovations from their means.

There errors ϵ_t are, by construction, uncorrelated with the factors

$$E(\epsilon_i) = 0; \quad E(\epsilon_i \tilde{f}_k) = 0$$

The key assumption in the model which is very important in order for the model is to be of value is that the errors are uncorrelated with each other.

$$E(\epsilon_i \epsilon_j) = 0$$

Although in more general models some limited amount of correlation between the residuals is permitted as Chamberlain & Rothschild (1983) demonstrate.

The pricing equation which relates the price of the expected returns in terms of the factor premia, based on the law of one price, is given

$$E(r) = \mathbf{1} r_f + \beta' \mathbf{\Lambda} \tag{3.26}$$

where r_f is the risk-free rate, $\mathbf{1}$ is a vector of one's, β is the matrix of beta coefficients, $\mathbf{\Lambda}$ is the vector of factor premia.

The pricing equation defines expected returns as a linear combination of the factor premia Λ, which are related to the prices of the factors. As we will be investigating whether aggregate volatility risk is priced, the Λ premia would need to be statistically significant. For robustness we will test this both on the whole sample as well as on sub-samples in a quarterly rolling regression. Significance does not however necessarily imply that aggregate volatility is priced since there could be a correlation between aggregate volatility and idiosyncratic volatility from the time series regression 3.25. We do not show the additional equations here as the $\mathbf{\Lambda}$ term contains all the chosen factors, but in order to control for idiosyncratic volatility we perform additional regressions of 3.26 with standard deviation of residuals from equation 3.25 included as an additional factor.

Chapter 4

Data & Methodology

As the thesis is organized into two main investigations, there are also two sets of source data used which are each best suited to facilitate the analysis. And from this source data returns and other higher moments are derived. In estimating volatility the source data used comprises high frequency price data for 29 US listed large stocks for the period starting July 2005 until December 2015 and therefore incorporates also the period of the great financial crisis of 2008. For the estimation of factor pricing in the Fama-French framework the main sources are the daily price data for the S&P 500 stocks as well as the Fama-French factors. The period is also chosen to coincide with the period for the high-frequency data in order to facilitate some comparisons i.e. from July 2005 until December 2015. Below follows now a more detailed description of the data and the preprocessing done to prepare the data for use.

4.1 High Frequency Data

Our high frequency (1-minute) price data incorporates 29 US-listed stocks shown in table 4.1 from which the realized risk measures are computed as well as forming the industry portfolios. Since the data had already been organized in 1-minute intervals, the initial data cleansing could be skipped, however the data still need to be converted into daily data as the analysis is conducted on daily volatility measures. Additionally, the high-frequency data in its raw form does not have any adjustments for dividends and stock splits which results in unnecessary jumps in price, so a key step in the data preparation was collecting and updating dividend and split data into the raw price data. The dividend and split data were obtained from Yahoo finance. Below are the summary statistics for the price data after adjusting for dividends and splits.

Table 4.1: Descriptive Statistics - Stock Prices

Ticker	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
GE	5.33	15.6	22.5	20.7	25.1	31.5
AAPL	5.05	17.1	40.0	48.8	74.7	131.0
AMZN	26.00	72.0	156.0	185.0	272.0	694.0
BAC	3.06	12.1	15.8	20.9	34.1	46.1
C	10.10	39.5	49.9	78.3	116.0	235.0
CMCSA	10.50	17.6	23.0	29.4	40.5	64.4
CSCO	12.10	17.0	20.3	20.4	23.5	30.2
CVX	38.80	57.0	72.3	77.2	98.4	127.0
DIS	14.20	27.4	31.6	45.4	59.7	121.0
HD	15.00	26.3	31.7	47.1	70.4	134.0
IBM	61.10	89.0	121.0	128.0	170.0	201.0
INTC	9.68	16.1	18.7	20.2	23.0	36.6
JNJ	38.50	50.2	54.8	63.6	78.9	105.0
JPM	13.90	33.6	38.0	40.9	47.5	69.6
KO	16.40	21.5	26.9	28.5	36.2	43.8
MCD	20.30	43.9	63.5	63.8	88.9	120.0
MRK	16.10	27.0	31.0	35.2	43.0	61.0
MSFT	12.70	21.8	24.3	27.5	29.9	56.6
ORCL	11.20	18.0	24.9	25.8	32.5	45.5
PEP	37.50	50.0	56.2	61.8	74.5	102.0
PFE	9.05	14.7	17.4	19.9	26.3	35.6
PG	35.70	48.9	54.6	58.2	71.0	90.5
QCOM	25.90	36.3	43.9	48.3	59.9	79.0
SLB	25.90	55.3	68.9	68.0	81.5	113.0
T	10.80	19.0	24.2	24.4	30.8	35.2
VZ	16.20	21.2	27.3	30.6	42.9	49.3
WFC	7.06	24.2	27.7	31.5	36.7	57.7
WMT	34.00	40.9	47.4	53.0	68.0	88.0
XOM	43.20	58.3	70.5	69.8	81.3	99.5

The analysis in terms of estimating volatility is carried out both on individual as well as portfolio level, thus we construct portfolio based on industry affiliation we use daily returns data to construct portfolio returns. Due to limited number of stocks for which the high-frequency data is available the portfolios do not completely reflect the common groupings as reflected by the various stock indexes, but the best attempt is made to match as closely as possible. The portfolios were formed by grouping the stocks based on industry affiliation as shown in Appendix A.1.

We see from the table 4.2 that with the exception of the bank portfolio, the median daily returns for the portfolios are quite comparable in magnitude with one another over the period. The median daily return for the bank portfolio

Table 4.2: Descriptive Statistics - Portfolio Returns

Portfolio	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
all	-0.0892	-0.00449	6.82e-04	3.51e-04	0.00590	0.114
banks	-0.2660	-0.00890	4.11e-05	4.26e-05	0.00961	0.246
tech	-0.0878	-0.00581	6.96e-04	5.47e-04	0.00765	0.111
oil	-0.1520	-0.00725	6.33e-04	2.63e-04	0.00824	0.178
cons	-0.0639	-0.00398	5.06e-04	4.09e-04	0.00520	0.097
indu	-0.3200	-0.00795	5.84e-04	1.92e-04	0.00923	0.347
coms	-0.1070	-0.00591	6.51e-04	3.53e-04	0.00689	0.167

is an order of magnitude lower than for the others likely reflecting the effect of the financial crisis of 2009. The mean return for the bank portfolio is also lower than for the other portfolios albeit by a slightly smaller margin. In terms of the extreme values we see that banks showing the most extreme minimum and maximum returns of -27% and 25% respectively. Interestingly, the industrials portfolio is showing even more extreme values than the bank portfolio.

In addition to the returns we also construct realized volatility measures used for estimating the underlying true volatility namely, realized volatility, bipower variance, jumps, realized negative semi-variance and realized positive semi-variance, using in this case only the overall portfolio.

Table 4.3: Descriptive Statistics for Volatility Measures

	Summary Statistics						Correlation				
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	rvol	rbpv	rjmp	rpsv	rnsv
rvol	5.58e-06	2.99e-05	5.63e-05	1.59e-04	1.29e-04	0.00988	1.000	0.927	0.799	0.886	0.886
rbpv	4.84e-06	2.27e-05	3.89e-05	1.06e-04	8.50e-05	0.00841	0.927	1.000	0.516	0.776	0.866
rjmp	0.00e+00	2.36e-06	9.93e-06	5.28e-05	3.60e-05	0.00431	0.799	0.516	1.000	0.781	0.634
rpsv	2.40e-06	1.36e-05	2.54e-05	7.94e-05	6.02e-05	0.00595	0.886	0.776	0.781	1.000	0.569
rnsv	2.32e-06	1.30e-05	2.52e-05	7.93e-05	5.98e-05	0.00639	0.886	0.866	0.634	0.569	1.000

Table 4.3 shows the summary statistics for the different variance estimators as well as the correlation between them. Here the mean variances are comparable to each other since by construction they should sum up to the realized variance (rvol). That is bipower variance plus jump variance as well as the semi-variances sum up to realized variance. In terms of the we see a strong correlation between realized volatility and the other variances as expected. Jump variance also shows a high correlation (80%) with realized variance as expected, however it shows a comparatively low correlation to bipower variance at just (52%). It also has a significantly higher correlation to positive semi-variance (78%) relative to negative semi-variance (63%).

4.2 Cross-Section Data

For the cross-section analysis we use the list of stocks in the S&P500 index taking only the stocks that are currently listed. Daily price data is taken over the period from 2005 until 2015 (i.e. same time-span as for the high frequency analysis) from which daily stock returns are calculated. Using this expanded list of stocks makes it possible to be able to split the sample into different risk categories and still have a diversity of stocks in each category. The main disadvantage with this approach is that this data as we do not have access to high frequency for the enlarged sample we're no longer able to construct realized variance measures that can then be used in the analysis. We will instead use the CBOE volatility index (VIX) as a proxy for market volatility. Whilst this may have some impact we believe that it should not affect the overall results obtained as the VIX is just another estimator for the underlying volatility which is also quite broadly used in the market.

For each of the S&P500 stocks we have constructed

Table 4.4: Descriptive Statistics - Market Data

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
MktExcRf	-8.95e-02	-4.70e-03	8.00e-04	3.27e-04	5.90e-03	1.13e-01
SMB	-3.78e-02	-3.30e-03	1.00e-04	3.83e-05	3.30e-03	3.85e-02
HML	-4.22e-02	-2.80e-03	-2.00e-04	-3.29e-05	2.40e-03	4.80e-02
RF	0.00e+00	0.00e+00	0.00e+00	4.87e-05	9.00e-05	2.20e-04
AggVol	9.89e+00	1.36e+01	1.71e+01	2.00e+01	2.31e+01	8.09e+01
AggVolInn	-1.02e+01	-6.47e+00	-2.93e+00	1.69e-15	3.10e+00	6.08e+01
AggVolIShort	-5.94e+00	-2.61e-01	-5.62e-03	-6.00e-19	2.39e-01	7.09e+00
AggVolIMed	-1.36e+01	-7.85e-01	-1.18e-01	5.73e-18	6.16e-01	1.34e+01
AggVolILong	-9.72e+00	-6.17e+00	-2.94e+00	1.70e-15	3.29e+00	4.49e+01

For the regression factors we have downloaded daily data of the Fama-French factors, namely: excess return on the market, SMB and HML returns. Summary statistics of the input market variables used for the Fama-French factors is shown in table 4.4. The Fama/French factors are constructed using combinations of the 6 value-weighted portfolios formed on size (small/large) and book-to-market (value/neutral/growth).

> SMB (Small Minus Big) is the average return on the three small portfolios minus the average return on the three big portfolios:

$$SMB = \frac{1}{3}(SmallValue + SmallNeutral + SmallGrowth) -$$

$$\frac{1}{3}(BigValue + BigNeutral + BigGrowth)$$

HML (High Minus Low) is the average return on the two value portfolios minus the average return on the two growth portfolios:

$$HML = \frac{1}{2}(SmallValue + BigValue) -$$
$$\frac{1}{2}(SmallGrowth + BigGrowth)$$

MktExcRf is the excess return of the market on the risk-free rate taking the value-weighted returns of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ

$$MktExcRf = R_m - R_f$$

RF is the simple risk-free rate equivalent to a 1-month Treasury bill rate

AggVol is the aggregate market volatility proxied by the CBOOE VIX index

AggVolInn is the daily innovation of the aggregate market volatility from its long term average

AggVolShort, AggVolMed and AggVolLong are the short, medium and long horizon wavelet decomposition of the aggregate volatility innovation. AggVolShort corresponds one-to-one to the smallest scale (i.e. highest frequency) which corresponds to the 2-4 day horizon. AggVolMedium is constructed by combining the scales 2-4 so that it has a horizon of 4-32 days. And finally the AggVolLong variable consists of all the rest of the decomposed signal.

Table 4.5: Correlation - Market Data

	MktExcRf	SMB	HML	AggVolInn	AggVolIShort	AggVolIMed	AggVolILong
MktExcRf	1.00000	0.16808	0.40368	-0.12606	-0.61933	-0.20890	-0.04419
SMB	0.16808	1.00000	-0.09450	-0.02382	-0.01125	-0.04237	-0.01597
HML	0.40368	-0.09450	1.00000	-0.07050	-0.21600	-0.09653	-0.03843
AggVolInn	-0.12606	-0.02382	-0.07050	1.00000	0.10557	0.22727	0.97928
AggVolIShort	-0.61933	-0.01125	-0.21600	0.10557	1.00000	0.17399	0.00011
AggVolIMed	-0.20890	-0.04237	-0.09653	0.22727	0.17399	1.00000	0.03962
AggVolILong	-0.04419	-0.01597	-0.03843	0.97928	0.00011	0.03962	1.00000

The correlation matrix for the market variables is shown in tn table 4.5. Here we observe that the decomposed short horizon aggregate volatility innovation, has a relatively strong correlation with excess market returns (62%)

especially when compared to the correlation between the short horizon and aggregate volatility innovation ((11%). The medium horizon volatility has a slightly higher correlation with aggregate volatility (23%) but the most correlated is the long horizon volatility (98%).

4.3 HAR Model

Financial data have been observed in many empirical studies to exhibit long-memory dependence particularly for volatility. A number of parametric methods have been proposed to deal with this persistence including ARCH/GARCH and stochastic family of models. In this paper we use instead the non-parametric HAR model proposed by Corsi (2009) in which the realized volatility is modeled as a linear function of lagged squared returns over different time periods to capture the long-memory features. Specifically we will use daily, weekly and monthly horizons as follows

$$RV_t = \alpha + \beta^{(D)} RV_{t-1} + \beta^{(W)} RV_{t-1}^{(5)} + \beta^{(M)} RV_{t-1}^{(22)} + \epsilon_t \tag{4.1}$$

where $\beta^{(D)}$, $\beta^{(W)}$, and $\beta^{(M)}$ are the daily, weekly and monthly betas. The RV, $RV^{(5)}$, and $RV^{(22)}$ correspond to the daily, weekly and monthly normalized variances. With the normalized multi-period volatilities computed as

$$RV_{t,t+h} = \frac{1}{h}(RV_{t+1} + RV_{t+2} + \cdots + RV_{t+h})$$

where $h = 1, 5, 22$

As we compare various HAR models in the paper, the right-hand side of equation 4.1 will vary depending on the given variance estimator that is being being modeled namely BPV, SV and additionally with or without jumps, namely:

$$RV_t = \alpha + \beta^{(D)} BPV_{t-1} + \beta^{(W)} BPV_{t-1}^{(5)} + \beta^{(M)} BPV_{t-1}^{(22)} + \epsilon_t \tag{4.2}$$

and similarly for the remaining estimators (BPV with jumps, SV, SV with jumps, SV with signed jumps).

4.4 Wavelet Scale Decomposition

Here we describe the wavelet details for the wavelet transformations being performed relating to the choice of wavelet and scale level. We also explain the approach how the wavelet decompositions are used in various scenarios and what questions they might help answer in each.

The choice of scale is on the one hand constrained by the length of the data series and on the other hand needs to be selected with attention to the underlying features of the data on which we'd like to gain further insight. In this case we have selected a scale of 10 which corresponds to 1024 (or approximately 4 years). This time horizon is appropriate as it covers an entire market cycle (which is approximately 4-6 years). This choice helps to avoid that the annual seasonal features which have been observed in financial data affect the smooth function. It also is longer than the 4-year presidential cycle in the US which has been shown in some studies to have an influence on stock returns so that we can isolate any cyclical patterns in the data from the smooth functions in order to draw out the underlying trends in the data.

For choice of wavelet we use the Daubechies family which is a family of compactly supported wavelet filters of various designed with a number of features that are particularly useful, especially for financial data series due to their smoothness and allowing the most accurate alignment in time between wavelet coefficients at various scales and the original time series.

There are three primary ways we use the wavelet scale decompositions in this paper for purposes of understanding the risk at different horizons as well as comparing the performance of the volatility estimators. In terms of understanding the horizon risks we investigate the aggregated level of risk obtained by a cumulation of the risks at higher scales up to a given scale. As an example the aggregated volatility at fifth scale would include the volatility of the smooth plus the volatility of the scales $10, 9, \ldots, 5$. This would then correspond to the volatility relevant to an investor that did not care about the short-term fluctuations as represented by the scales 1 to 4.

A second way we use the wavelet scale decomposition is to investigate the scale-specific features that are relevant to risk and how they compare across scales. In this usage we can answer the question whether the risk is significant at lower or higher scales and also whether the risk is increasing or decreasing as the scales increase. Additionally we can answer the question whether the model a good fit at the lower scales or higher scales and whether that fit is increasing or decreasing.

The third and final way we use scale decomposition is to answer the question about the performance stability of the model between in-sample estimation and out-of-sample forecasts at different scales. As we're interested not just in modeling volatility but also being able to forecast accurately then this test is quite important as well to ensure we did not have a spurious regression result and can assure similar performance across different samples.

4.5 Cross-Sectional Analysis

In the cross-section analysis we're interested primarily in two questions. The first question concerns whether aggregate volatility priced in the cross-section of returns. The second questions concerns whether stocks with a high sensitivity to aggregate volatility earn lower subsequent returns and higher contemporaneous average returns.

In terms of investigating whether aggregate volatility is priced we follow the standard practice of running a time series regression to determine the beta coefficients for the systemic factors and subsequently running a cross-sectional regression of the beta sensitivities on expected returns to determine the factor premia. As we already investigated the topic of risk at different horizons in the section 5 on volatility, we focus in this chapter 6 investigating the pricing of risk at different horizons.

To determine the factor premium we perform the analysis in these steps:

- Time series regression for each of the 433 stocks to determine the betas
- Regression of the estimated betas on the average returns to determine the factor premia

As factors we use the Fama-French factors plus aggregate volatility innovation. As a robustness check we also regress the short, medium and long-term aggregate volatility innovations in conjunction with the Fama-French factors. The betas obtained are then used in the factor pricing regressions. To control for idiosyncratic in the factor pricing regression, we also include the average idiosyncratic volatility in this regression.

We perform the exact same regressions also on the wavelet decomposed series to obtain the factor pricing on a scale-by-scale basis. With these factor premia we're then able to make some analysis regarding the pricing of aggregate volatility risk at different horizons. The two questions that we're able to answer which whether aggregate volatility premium is statistically significant at each horizon and secondly whether the factor pricing becomes more important as the scale increases.

Finally we perform rolling regressions on a quarterly basis on both the normal time series data as well as the wavelet scale decomposed series and estimate factor pricing using a rolling 2-year window for the regressions. We then plot the estimated factor pricing for aggregate volatility along with confidence bands to see whether it's significant and its evolution through time.

The second main question for this section is whether stocks which have a high sensitivity to aggregate volatility earn lower future average returns and higher contemporaneous average returns. To answer this question we compare the future average returns from five sorted portfolios that are sorted according to size of the beta coefficient for aggregate volatility factor from the time series regression. Portfolios are then sorted from highest to lowest beta size and within each portfolio we compute equal-value weighted returns for current and next quarter. And finally we compare the returns from the two portfolios with the highest and lowest sensitivity to aggregate volatility. As a robustness check we perform the regression both for the entire sample as well as on a rolling quarterly basis.

Chapter 5

Empirical Results - Volatility

As described in the data section, for the volatility analysis we use the realized data for the 29 stocks to perform the analysis on different levels of aggregation, namely:

- Overall market aggregation level where all 29 stocks are combined into a single "market" portfolio in order to be able to gain some understanding into total market behavior

- Industry level whereby stocks within the same or closely related industries are grouped together

Whilst an analysis was also carried out on individual stock level during the process, discussion of the results on an individual stock level is out of scope for this thesis. In scope for the analysis is to understand the structure of volatility and how it evolves over time and how the information helps improve estimation and forecasting performance. The analysis is divided into firstly the variance decomposition into scales for the various realized measures and their evolution through time. Secondly, we investigate whether there is an improvement in the parameter estimations between the normal time series regressions vs. regression on different scales. And we also look at whether there is a difference in regression performance on a scale level between the aggregate market and industry portfolios. Portfolios should have theoretically smaller idiosyncratic volatility whilst the market portfolio should have essentially no idiosyncratic volatility so this final test could in fact be seen as a test for whether idiosyncratic volatility has an impact on parameter estimations on different levels of aggregation.

5.1 Structure of Variance Decomposition

For each of the data series i.e. market portfolio and industry portfolios we constructed we analyze the distribution of variance decomposition. For robustness we will compare the overall distributions for three samples to be able to compare between test and training samples namely:

- Estimating across entire data sample from July 2005 until December 2015 as a baseline for comparison
- Training sample which runs from start of time series until December 2014
- Test sample of one year from January 2015 until end of data series on December 2015 (i.e. 250 forecast data points)

Looking at the energy distribution across scales for the full sample data in table 5.1 we see that there is more energy concentrated in the lower scales (shorter range) than in the longer-range scales. Furthermore we note that the different measures of volatility contain roughly similar amounts of energy at each of the scales.

However, there is a fine distinction which can be observed in that the energy distributions for realized volatility and bipower variation are very similar to each other across different scales, and similarly for the energy distribution between the positive and negative semi-variances but there are sizable differences between the two groups across scales. For example for the shortest scale "D1", realized volatility and bipower variance contain 20.5 and 19.5 of the total respectively whereas the positive and negative semi-variances contain 24.1 and 26.9 respectively. This is a reflection of the fact that the semi-variance measure are more finely tuned to the daily change in the sign of returns and therefore more of the high frequency information content.

Furthermore, a comparison between the two semi-variances shows the negative semi-variance has more of its energy concentrated in the shorter range scales than the positive semi-variance. This is also in line with economic intuition as the negative returns are associated with greater fear among investors and greater short-term trading to protect investments.

For robustness we also evaluate the energy distribution with different samples and sample sizes by breaking the data into training and a smaller testing

Table 5.1: Variance Wavelet Decompostion - Full Sample

Type	Portfolio	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10	s10
Realized Variance	all	20.5	11.3	9.1	6.9	5.4	8.1	8.1	9.9	9.0	6.2	5.5
	banks	23.3	14.9	8.9	6.9	4.9	4.9	5.0	8.4	10.1	6.8	6.0
	tech	19.0	11.9	9.1	6.5	6.6	9.5	10.3	8.7	7.1	5.7	5.6
	oil	24.6	12.6	7.5	6.9	7.5	10.6	10.0	7.7	5.8	3.5	3.3
	cons	31.2	15.5	10.6	6.0	5.2	6.4	7.1	6.6	5.2	3.5	2.6
	indu	48.1	24.2	12.2	6.6	1.6	2.2	1.8	1.4	1.1	0.4	0.5
	coms	31.4	15.1	9.6	5.7	4.4	6.9	7.5	6.9	5.1	3.5	4.0
BiPower Variance	all	19.5	10.6	7.3	7.0	6.5	8.5	8.9	10.5	9.2	6.4	5.5
	banks	12.0	10.0	9.0	8.6	6.8	5.1	4.8	10.8	13.6	10.5	8.7
	tech	13.6	8.1	5.7	6.7	7.9	10.9	11.8	11.3	9.5	7.3	7.1
	oil	23.9	11.9	6.1	7.5	8.1	11.0	10.4	8.1	5.8	3.7	3.5
	cons	31.7	15.7	9.8	6.1	5.2	6.3	7.2	6.6	5.2	3.5	2.6
	indu	16.1	9.5	6.7	6.5	6.4	9.2	7.9	11.1	10.8	7.9	7.9
	coms	22.1	11.2	7.6	6.0	5.6	9.1	10.2	9.6	7.8	5.5	5.2
Jump Variance	all	39.4	19.4	11.0	5.9	3.3	4.2	3.7	4.1	4.0	2.6	2.5
	banks	41.9	23.6	10.1	4.3	2.7	3.3	3.3	3.4	3.7	2.0	1.8
	tech	42.1	21.6	12.7	6.1	3.9	3.7	3.6	2.1	1.6	1.3	1.3
	oil	40.3	18.7	11.6	6.7	5.0	4.9	4.2	3.1	2.7	1.5	1.3
	cons	44.0	18.5	11.4	4.9	4.1	4.0	3.7	3.5	2.7	1.8	1.5
	indu	50.0	25.0	12.5	6.6	1.3	1.8	1.4	0.8	0.4	0.1	0.1
	coms	48.4	23.6	12.2	6.4	2.9	1.7	1.8	1.3	0.6	0.4	0.7
Signed Jump	all	43.5	26.3	14.2	6.9	3.8	2.7	1.7	0.4	0.2	0.3	0.1
	banks	48.4	26.9	12.7	3.4	2.5	2.9	2.1	0.4	0.4	0.1	0.1
	tech	44.3	27.2	14.4	6.3	3.8	2.4	1.0	0.4	0.3	0.0	0.0
	oil	48.0	22.7	12.5	9.0	4.4	1.8	0.9	0.4	0.1	0.0	0.0
	cons	45.8	24.8	14.2	7.8	3.7	1.8	0.9	0.6	0.3	0.2	0.0
	indu	49.9	25.0	12.5	5.9	4.9	1.3	0.3	0.0	0.0	0.0	0.0
	coms	49.2	24.8	11.9	6.5	4.0	2.4	0.8	0.2	0.1	0.1	0.0
Positive Semi-Variance	all	24.1	14.3	10.0	6.4	4.6	7.4	7.7	7.6	7.6	5.5	4.9
	banks	36.8	19.9	10.0	4.5	3.5	4.1	3.8	4.4	5.8	3.7	3.4
	tech	24.1	15.0	8.9	5.2	5.7	8.6	8.7	7.4	6.4	5.0	5.0
	oil	28.1	11.4	6.7	5.8	7.2	11.1	10.2	7.4	5.5	3.4	3.2
	cons	27.5	14.3	10.5	5.4	5.1	7.3	8.5	7.4	6.0	4.4	3.7
	indu	48.8	24.5	12.3	6.3	3.2	1.8	1.1	0.8	0.7	0.2	0.3
	coms	35.1	17.0	9.9	6.4	4.8	5.6	5.8	5.1	3.9	3.0	3.4

Table 5.2: Variance Full Decomposition Comparison - Portfolio(All)

Type	**Portfolio**	**d1**	**d2**	**d3**	**d4**	**d5**	**d6**	**d7**	**d8**	**d9**	**d10**	**s10**	**waveVar**	**sigma.sq**	**Bias**
Full Sample	Realized Vol.	20.5	11.3	9.1	6.9	5.4	8.1	8.1	9.9	9.0	6.2	5.5	2.00e-07	2.00e-07	0
	BiPower	19.5	10.6	7.3	7.0	6.5	8.5	8.9	10.5	9.2	6.4	5.5	1.00e-07	1.00e-07	0
	PosSV	24.1	14.3	10.0	6.4	4.6	7.4	7.7	7.6	7.6	5.5	4.9	1.00e-07	1.00e-07	0
	NegSV	26.9	14.8	10.4	7.4	5.6	6.5	5.8	8.0	6.5	4.3	3.8	1.00e-07	1.00e-07	0
Training Sample	Realized Vol.	18.1	9.8	8.3	6.6	5.4	8.7	8.9	11.0	10.0	7.3	6.1	2.00e-07	2.00e-07	0
	BiPower	17.7	9.7	6.8	6.8	6.6	9.0	9.5	11.3	9.8	7.1	5.8	1.00e-07	1.00e-07	0
	PosSV	24.1	14.3	10.0	6.3	4.6	7.4	7.7	7.7	7.7	5.7	4.6	1.00e-07	1.00e-07	0
	NegSV	21.8	10.9	8.4	6.9	5.5	7.8	7.4	10.8	8.8	6.2	5.5	0.00e+00	0.00e+00	0
Testing Sample	Realized Vol.	39.2	22.9	15.4	9.3	6.0	3.9	2.2	1.1*				2e-07	2e-07	0
	BiPower	42.6	22.4	14.4	8.9	5.3	3.6	1.9	0.9*				1e-07	1e-07	0
	PosSV	35.4	15.3	15.4	11.1	7.1	7.8	5.0	2.8*				0e+00	0e+00	0
	NegSV	39.6	24.5	15.4	8.7	5.7	3.3	1.8	0.9*				2e-07	2e-07	0

sample as well as the full sample data. In order to compare like with like however, we have to restrict this comparison analysis to only seven scales to ensure that all samples are decomposed in the same number of scales. This is necessary since the testing sample has by convention a much smaller number of data points as compared to the training dataset which in this case allows decomposition of the testing sample into a maximum of seven scales (i.e. 250 data points in one year). Therefore, to facilitate a comparison of the wavelet decompositions for the training, testing and full sample will be restricted to seven scales. However, for the rest of the analysis in this paper the training sample will be decomposed to ten scales and testing sample to seven in order to utilize as much information present in the signal as possible.

For the training sample data we see that the distribution of the variance decomposition is approximately in line with that from the full sample. This makes sense as the sample periods are nearly identical differing only by the difference of 1 year. The negative semi-variance energy shows some small decline in the training sample for the lower scales (i.e. shorter range) and an increase in the energy content of the higher (longer range) scales. On the lowest scale "D1" the energy content declined from 26.9% to 21.8% and for the highest scale "D10" & "S10" rose from 4.3% & 3.8% to 6.2% & 5.5% respectively.

When comparing to the distributions from the testing sample however we see some important differences from the training and full samples. The energy is much more intensely focused in the higher frequency scales. For each of the first three scales the energy content has doubled as compared to that of the training and full samples. And correspondingly the energy content at the higher scales has dropped sharply. A possible explanation could be that there was some kind of event in 2015 which caused markets to become more volatile than over the preceding period. Alternatively it could be as a result of the smaller sample size due to the averaging effect when analyzing the longer time-frame. Nevertheless this finding implies different distributions for the volatility between the training and testing periods and so could affect the accuracy of the forecasting later.

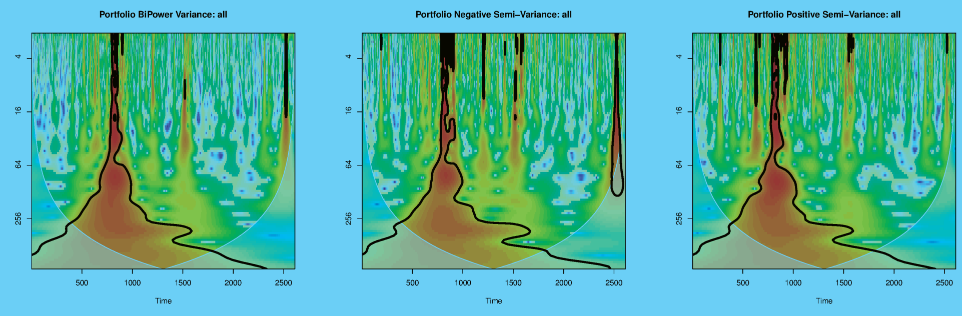
5.2 Evolution of Variance Decomposition

Analyzing the variance decomposition for each of the time series can provide a perspective on how much of the signal power is contained on each of the different scales and therefore can give a good indication on the importance of variability at each of the scales to the overall variability of the series. And by performing a rolling analysis we can get an understanding if there are changes in the power distribution over time, for example during upswings and downswings or during times of financial crises.

In doing a rolling analysis the challenge is always to decide on an appropriate window length. A longer window length is statistically advantageous as it means that there are more data points that are feeding into the model and so should be helpful in providing better estimates. On the other hand from a financial perspective there are two perspectives which favor restricting the amount of data used. Firstly, from a valuation perspective the most recent data is the most relevant for both current and future valuation and therefore should have more weight. Secondly, over time the economic and financial environment as well as the company fundamentals will change so that old information may no longer be applicable in current circumstances.

For this analysis we will use a window period of two years with daily data. The variance decomposition evolution is calculated on a rolling basis using the maximum discrete overlap method which is useful for data series that are of length not a power of two and because it is shift invariant. For comparison we also use the continuous wavelet transform which is shown in the figure 5.1. This method has the advantage that it provides a good overview of the energy distribution, even though it has the slight disadvantage compared to the discrete method that it does not present a sufficiently clear energy breakdown by scale.

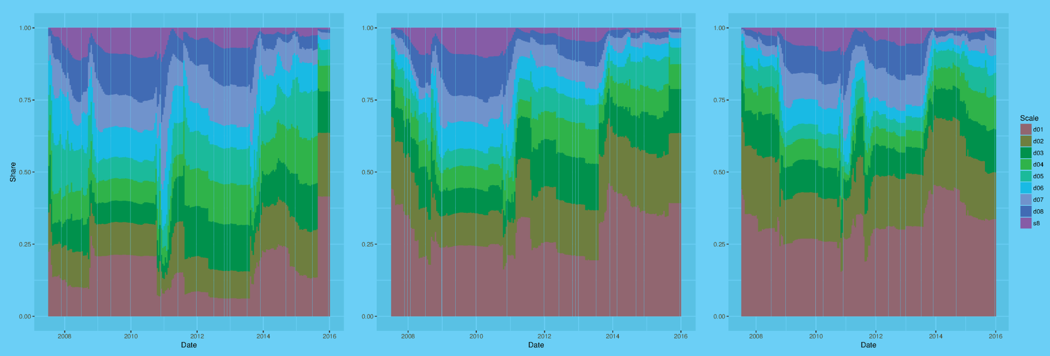
We observe in the figure 5.1 that the energy distribution is not constant throughout the period. For most of the time the signal energy is mainly concentrated in lower frequency components (i.e. higher scales). However during certain key periods (corresponding approximately to 2009, 2011 and 2015) for all three realized volatility estimators the energy content in the higher frequency part (i.e. shorter scales) is quite prominent and statistically significant. As these periods have been associated with extreme market volatility this sug-



(a) BiPower Variance (b) Neg. Semi-Variance (c) Pos. Semi-Variance

Figure 5.1: Wavelet Decomposition

gest that during periods of extreme market volatility then volatility tends to be more dominated by short-term components than is the case during more normal market periods. An interpretation of this figure is that is more important at higher scales than the lower scales. However during times of heightened market uncertainty then volatility becomes even more important at lower scales compared to higher scales. Informally one could interpret this chart to say that at the lower scales (the short-term horizons) volatility is generally not important until there is some sort of crisis and then it is extremely important.



(a) BiPower Variance (b) Neg. Semi-Variance (c) Pos. Semi-Variance

Figure 5.2: Rolling Variance Decomposition

To show the exact energy distribution by scale the data series has been decomposed using the maximum overlap method to 10 scales. The figure 5.2 shows the rolling variance decomposition for only three of the volatility estimators used in this paper. Additional charts can be found in the appendix.

Beginning with the bipower variance rolling decomposition, the energy concentration is highest (average 21%) in the shortest scale (D1) and the rest of the energy is about evenly distributed between the rest of the scales (averaging between 6% - 11%). There are however several periods where the energy content in the lowest scales spikes up significantly so that much more of the variability in the volatility is due to short term innovations. These timing for these periods coincides with similar observations in the wavelet transform figure 5.1. This also happens to corresponding quite well with periods where the market was experiencing elevated turmoil during and after the great financial crisis (i.e. 2009, 2011, 2013 and 2015).

For both positive and negative semi-variances the energy content of the signal is much more concentrated in the shorter scales. For these we also observe spikes in the energy content of the shorter scales during times of elevated market turmoil. For the period after 2014 the negative semi-variance shows an elevated energy content in the shorter scales for longer than the other two variances, suggestive of the presence of stronger persistence in the signal, this topic however is beyond the scope of this paper and is not investigated further.

A number of conclusions could be drawn from these observations. Firstly, compared with the bipower variation the semi-variances have a higher concentration of their variability in the shorter time-frames and therefore leaving these out would result in greater inaccuracy in predicting volatility as compared with the bipower variation. So for an investor looking to focus on longer horizon strategy they would be better off using bipower variation as a proxy for volatility as opposed to using the semi-variances. Secondly, periods of market volatility cause the energy content of the shorter horizon scales to increase further. On the whole however one can see that during times of market turmoil then the short-term volatility component becomes more important overall so the uncertainty is much greater for investors with a short-horizon outlook.

5.3 Horizon Volatility

We have seen from the volatility decomposition plots and tables that the short horizon scales are the most important in terms of their contribution to the overall energy content of volatility. Additionally we also observed in the evolution of volatility that during periods of high uncertainty the contribution to the energy content from the lower scales in fact increases. So we might therefore be interested to find out whether the increase in contribution to the energy content is due purely to an increase in the variability of volatility, or if the level of volatility is also changing.

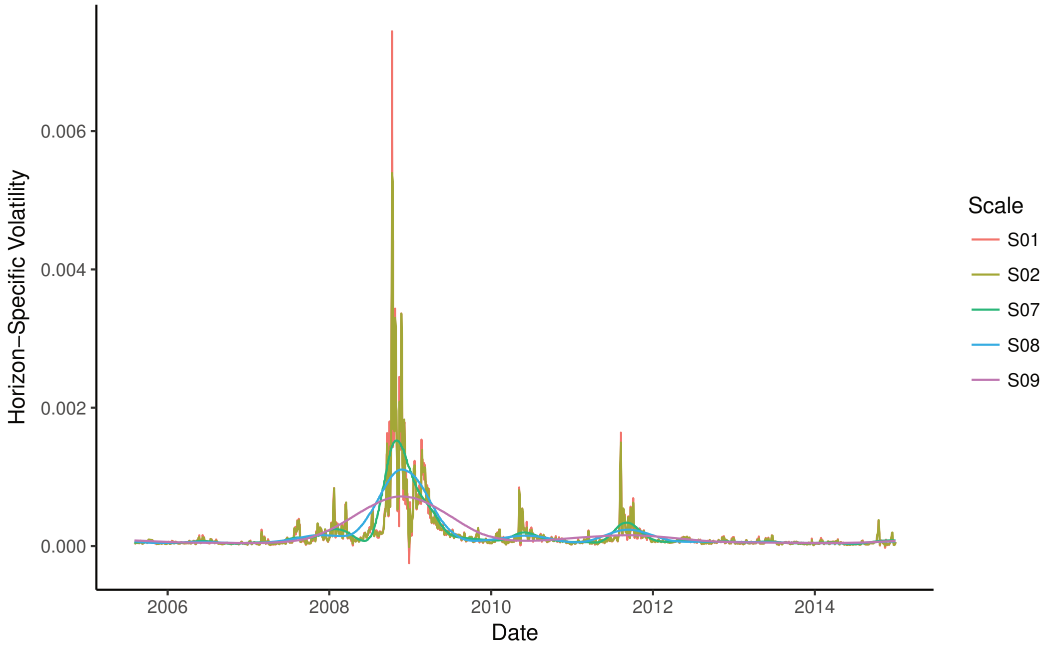


Figure 5.3: Horizon-Relevant Volatility (BPV w/ Jumps)

Figure 5.3 shows a chart of the scale decomposed volatility showing the volatility relevant for only the scales $1, 2, 7, 8 \,\&\, 9$. The scales correspond to the horizons $2-4$ days, $4-8$ days, $128-256$ days (approx. 6 - 12 months), $256-512$ days (approx. 1 - 2 years), and $512-1024$ days (approx. 2 - 4 years). What the chart shows is there is a big difference in the level of volatility across different horizons during certain key periods around 2009, 2011, and 2015. The volatility at the shortest horizon 2 - 4 days (S01) has a volatility more than 4 times the volatility at the 1 - 2 years (S08) horizon and over 3 times that at the 6 - 12 months (S07) horizon. During normal times the volatilities across the different horizons are pretty much similar, however during the key periods the divergence is quite steep. These periods happen to be times when there was stress in the markets so it is possible that the shorter horizon volatility is much more sensitive uncertainty in the market. Therefore when measured in terms of the level of volatility, we can say that short-term horizon investors face a significantly higher level of risk than longer horizon investors particularly if there would occur some disturbance in the markets.

5.4 Estimation

In terms of estimation we analyze and compare the various risk measures to see how they perform on each of the portfolios. In the first regressions we compare the performance of the various risk measure namely bipower, bipower with jumps, bipower with signed jumps, semi-variances and semi-variances with jumps to establish firstly how well they perform. To examine the relative stability of coefficients we compare the coefficients obtained from regressing the portfolios we've constructed made of stocks from similar industry groups. Whilst this analysis is somewhat limited due to the small number of stocks in the sample (i.e. 29) it nonetheless suggestive that there may be quite significant differences across portfolios regarding beta coefficients in terms of both economic and statistical significance as shown in the next set of regression tables.

For the rest of the tables in this section we're looking at three main comparisons. Firstly, to see the stability of the estimated volatility models we compare across sector to see if the regression statistics are stable. Secondly, we compare across samples (full, training and test samples) to see if the R^2 are stable. And finally we compare across scale to see at which horizons the model is still performing reasonably well and where it is no longer performing well perhaps due to lack of fit or perhaps a result of noise. This means that if some models would perform better on short or long horizon data then we could still analyze that.

Table 5.3: Bipower Volatility Regressions - Portfolio

	Dependent variable:						
	Realized Vol.						
	all	banks	tech	oil	cons	indu	coms
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
rbpv1	0.363***	0.770***	0.535***	0.033	0.091***	0.415*	0.121***
	(0.034)	(0.051)	(0.038)	(0.032)	(0.032)	(0.223)	(0.044)
rbpv5	0.695***	0.555***	0.624***	0.915***	0.566***	0.588	0.779***
	(0.058)	(0.080)	(0.059)	(0.055)	(0.061)	(0.361)	(0.076)
rbpv22	0.204***	0.257***	0.063	0.133***	0.416***	0.189	0.234***
	(0.050)	(0.072)	(0.048)	(0.049)	(0.060)	(0.305)	(0.068)
Constant	0.00003***	0.00003	0.00003***	0.00005***	0.00002***	0.0002*	0.00004***
	(0.00001)	(0.00004)	(0.00001)	(0.00001)	(0.00001)	(0.0001)	(0.00001)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.547	0.490	0.578	0.440	0.329	0.031	0.349
Note:							*p<0.1; **p<0.05; ***p<0.01

Bipower An evaluation of the performance of the bipower variance estimator across the different portfolios as shown in table 5.3. The adjusted R^2 are relatively high (in the range 35% - 55%) for all portfolios except for industrial portfolio regression which has an adjusted R^2 of only 3%. Judging in terms of statistical and economic significance of the coefficients, we say that except for industrials the model is a relatively good fit.

Table 5.4: Semi-Variances Regression - Portfolios

	Dependent variable: Realized Vol.						
	all	banks	tech	oil	cons	indu	coms
rpsv1	0.108^{***}	0.080^{***}	0.185^{***}	-0.201^{***}	0.221^{***}	0.005	0.079^{**}
	(0.038)	(0.027)	(0.038)	(0.061)	(0.065)	(0.035)	(0.041)
rpsv5	0.225^{***}	-0.122^{*}	0.442^{***}	0.535^{***}	0.288^{**}	-0.330^{***}	0.527^{***}
	(0.083)	(0.071)	(0.084)	(0.136)	(0.146)	(0.088)	(0.090)
rpsv22	0.263^{*}	0.439^{***}	-0.189	0.953^{***}	0.173	1.830^{***}	-0.022
	(0.139)	(0.102)	(0.133)	(0.246)	(0.322)	(0.168)	(0.146)
rnsv1	0.286^{***}	0.401^{***}	0.320^{***}	0.234^{***}	-0.111^{**}	-0.012	-0.011
	(0.048)	(0.047)	(0.038)	(0.059)	(0.050)	(0.030)	(0.037)
rnsv5	0.728^{***}	0.778^{***}	0.361^{***}	0.920^{***}	0.617^{***}	-0.217^{***}	0.304^{***}
	(0.100)	(0.097)	(0.081)	(0.128)	(0.123)	(0.078)	(0.088)
rnsv22	0.247	0.401^{***}	0.686^{***}	-0.730^{***}	0.515	0.054	0.824^{***}
	(0.159)	(0.138)	(0.151)	(0.260)	(0.318)	(0.135)	(0.143)
Constant	0.00001^{*}	0.0001	0.00002^{**}	0.00004^{***}	0.00001^{**}	0.0002^{**}	0.00003^{***}
	(0.00001)	(0.00005)	(0.00001)	(0.00001)	(0.00001)	(0.0001)	(0.00001)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.511	0.430	0.526	0.425	0.321	0.053	0.310
Note:					*p<0.1; **p<0.05; ***p<0.01		

Semi-variance Comparison of the regression results for the semi-variance estimator on industry portfolios is shown in table 5.4. The regression results for industrial portfolio is again fairly weak with an adjusted R^2 of only 5% (a marginal improvement from 3% with BPV regression). For the other portfolios the results show relatively good adjusted R^2 figures and statistically and economically significant coefficients across all portfolios.

BPV and Semi-Variances w/ Jumps Comparison of the regression results for the BPV with jumps and SV with jumps estimators are shown in tables 5.5 and 5.6. Including the jumps in the regression does not significantly improve the adjusted R^2 results for the portfolio. The adjusted R^2 improves only marginally for both the BPV and SV regressions with jumps. Additional

Table 5.5: Bipower Vol. w/ Jumps Regression - Portfolios

	Dependent variable:						
	Realized Vol.						
	all	banks	tech	oil	cons	indu	coms
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
rbpv1	0.381***	0.785***	0.549***	0.094***	0.210***	0.416*	0.132***
	(0.034)	(0.051)	(0.038)	(0.034)	(0.037)	(0.223)	(0.044)
rbpv5	0.741***	0.592***	0.640***	0.913***	0.543***	0.588	0.784***
	(0.058)	(0.081)	(0.059)	(0.055)	(0.061)	(0.361)	(0.076)
rbpv22	0.205***	0.260***	0.060	0.156***	0.438***	0.189	0.234***
	(0.050)	(0.072)	(0.048)	(0.049)	(0.059)	(0.305)	(0.068)
rjmp1	−0.178***	−0.082***	−0.085***	−0.335***	−0.492***	−0.004	−0.064**
	(0.037)	(0.024)	(0.032)	(0.068)	(0.077)	(0.021)	(0.028)
Constant	0.00003***	0.00003	0.00003***	0.0001***	0.00002***	0.0002*	0.00004***
	(0.00001)	(0.00004)	(0.00001)	(0.00001)	(0.00001)	(0.0001)	(0.00001)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.551	0.492	0.580	0.445	0.341	0.030	0.350
Note:							*p<0.1; **p<0.05; ***p<0.01

regression results including the signed jumps are presented in Appendix C.

In terms of R^2 the models seem to be performing consistently and producing quite similar R^2 across the different sectors. All the models struggle with the industrials sector showing very low R^2 values, however this could just be an issue with a small sample set. Worth noting as well in these regressions is the relative stability of the adjusted R^2 for the portfolio formed from all stocks (ranging from 51% - 55%) which is important because the conclusions of the APT are dependent on either the R^2 being high or the number of assets being large and this is also a motivation to conduct the analysis on the entire portfolio of stocks instead of separating into sectors.

In the next section we analyze the estimation performance at each scales separately, comparing between the various volatility estimators. Finally we obtain different horizon volatilities where the volatility at each horizon is composed of the detail component at the given scale plus the smooth (i.e. all the lower frequency components). The focus will be on the performance statistics for the estimation regression and prediction.

From the regression tables 5.7, 5.8, 5.9 and 5.10, there are some overall facts which seem to run across each of the regressions (both BPV and SV) as

Table 5.6: Semi-Variances w/ Jumps Regressions - Portfolio

	Dependent variable:						
	Realized Vol.						
	all	banks	tech	oil	cons	indu	coms
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
rpsv1	0.564***	0.936***	0.623***	0.046	0.707***	0.646***	0.403***
	(0.049)	(0.056)	(0.044)	(0.072)	(0.078)	(0.132)	(0.052)
rpsv5	0.207***	−0.161**	0.261***	0.586***	0.466***	−0.363***	0.395***
	(0.080)	(0.067)	(0.080)	(0.135)	(0.144)	(0.087)	(0.089)
rpsv22	0.043	0.329***	−0.188	0.692***	−0.401	1.644***	−0.012
	(0.135)	(0.096)	(0.126)	(0.248)	(0.319)	(0.171)	(0.143)
rnsv1	0.376***	0.867***	0.731***	0.272***	0.048	0.639***	0.254***
	(0.047)	(0.052)	(0.044)	(0.059)	(0.051)	(0.133)	(0.046)
rnsv5	0.703***	0.549***	0.279***	0.797***	0.317**	−0.207***	0.286***
	(0.097)	(0.093)	(0.077)	(0.129)	(0.124)	(0.078)	(0.086)
rnsv22	0.399***	0.279**	0.532***	−0.443*	0.992***	−0.099	0.615***
	(0.153)	(0.130)	(0.143)	(0.262)	(0.314)	(0.137)	(0.142)
rjmp1	−0.761***	−1.012***	−0.812***	−0.574***	−1.130***	−0.656***	−0.469***
	(0.056)	(0.059)	(0.049)	(0.093)	(0.107)	(0.130)	(0.049)
Constant	0.00002***	0.00002	0.00003***	0.00004***	0.00002***	0.0001	0.00003***
	(0.00001)	(0.00004)	(0.00001)	(0.00001)	(0.00001)	(0.0001)	(0.00001)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.547	0.493	0.576	0.434	0.352	0.063	0.335
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 5.7: Bipower Variance Scales Regression - Portfolio(All)

	Dependent variable: Realized Vol.										
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	S10
rbpv1	−0.381***	0.953***	1.625***	1.648***	1.202***	3.395***	3.771***	7.533***	23.982***	−45.193***	47.687***
	(0.038)	(0.024)	(0.015)	(0.018)	(0.029)	(0.127)	(0.337)	(0.519)	(1.152)	(4.269)	(3.090)
rbpv5	−1.392***	−0.453***	−0.960***	−0.416***	−0.022	−2.196***	−2.571***	−7.990***	−28.098***	57.218***	−59.332***
	(0.215)	(0.100)	(0.027)	(0.019)	(0.031)	(0.152)	(0.413)	(0.640)	(1.423)	(5.271)	(3.817)
rbpv22	−9.963***	−8.467***	−0.235*	0.341***	−0.075***	0.172***	0.114	1.865***	5.554***	−10.659***	13.071***
	(0.888)	(0.404)	(0.124)	(0.035)	(0.016)	(0.035)	(0.083)	(0.124)	(0.272)	(1.003)	(0.727)
Constant	−0.000	−0.000	0.000	−0.000	−0.000	−0.000	0.00000	−0.00000***	−0.00000***	−0.000	0.00001***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.330	0.414	0.844	0.919	0.946	0.984	0.984	0.998	0.999	0.999	1.000
Note:											*p<0.1; **p<0.05; ***p<0.01

well as across all scales. Firstly the coefficients are nearly all statistically and economically significant and secondly, the constant is approximately zero in all the regressions. This is a strong indication that the fit of the models to the data is very good.

Turning to adjusted R^2 measure we see relatively high values for the higher

Table 5.8: Semi-Variances Scales Regression - Portfolio(All)

	Dependent variable: Realized Vol.										
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	S10
rpsv1	-0.404^{***}	0.497^{***}	0.999^{***}	1.326^{***}	1.467^{***}	1.559^{***}	1.618^{***}	1.552^{***}	1.626^{***}	1.649^{***}	1.641^{***}
	(0.028)	(0.025)	(0.010)	(0.009)	(0.008)	(0.006)	(0.005)	(0.005)	(0.003)	(0.004)	(0.003)
rpsv5	-1.722^{***}	-0.491^{***}	-0.690^{***}	-0.513^{***}	-0.526^{***}	-0.579^{***}	-0.649^{***}	-0.566^{***}	-0.656^{***}	-0.685^{***}	-0.674^{***}
	(0.165)	(0.090)	(0.020)	(0.009)	(0.008)	(0.008)	(0.006)	(0.006)	(0.004)	(0.005)	(0.004)
rpsv22	-7.986^{***}	-7.832^{***}	-0.895^{***}	-0.220^{***}	0.029^{***}	0.013^{***}	0.031^{***}	0.014^{***}	0.029^{***}	0.035^{***}	0.033^{***}
	(0.730)	(0.405)	(0.118)	(0.023)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
rnsv1	-0.555^{***}	0.813^{***}	1.201^{***}	1.305^{***}	1.587^{***}	1.669^{***}	1.658^{***}	1.733^{***}	1.658^{***}	1.635^{***}	1.643^{***}
	(0.037)	(0.030)	(0.015)	(0.009)	(0.009)	(0.007)	(0.005)	(0.005)	(0.003)	(0.004)	(0.004)
rnsv5	-0.897^{***}	-0.915^{***}	-0.688^{***}	-0.452^{***}	-0.649^{***}	-0.720^{***}	-0.697^{***}	-0.787^{***}	-0.695^{***}	-0.667^{***}	-0.677^{***}
	(0.204)	(0.122)	(0.027)	(0.010)	(0.010)	(0.008)	(0.007)	(0.007)	(0.004)	(0.005)	(0.005)
rnsv22	-2.666^{***}	-7.487^{***}	0.073	-0.234^{***}	0.062^{***}	0.056^{***}	0.038^{***}	0.055^{***}	0.038^{***}	0.032^{***}	0.034^{***}
	(0.863)	(0.549)	(0.139)	(0.025)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	−0.000	−0.000	0.000	0.000	0.000	−0.000	0.000	0.000^{*}	−0.000	-0.000^{**}	0.000^{***}
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.620	0.473	0.930	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000
Note:											*p<0.1; **p<0.05; ***p<0.01

scales in the regressions of both BPV and SV volatility measures. The expectation is that the adjusted R^2 is rising as the time horizon increases since the short term fluctuations smooth out over the long term and this seems to be confirmed by the data. This increase in adjusted R^2 is also quite steep as the scales go up. We also observe that the jump variable has a different impact on different scales. This result confirms the prediction made in section 2.2 that the HAR model should show a much better fit at higher scales than at lower ones. It is also in line with the observation in 5.1 that the realized variance estimators are more important and statistically significant at the higher scales.

For bipower in table 5.9 the adjusted R^2 is already 84% by the third scale and over 90% for the higher scales, which gives some hints that forecasting could already be significantly improved by focusing on the 3rd and higher scales. The first and second scales are quite low by comparison with only 33% and 41% respectively. Addition of the jump variable however adds a huge improvement with the adjusted R^2 for the first scale improving to 58%. Adjusted R^2 for the second scale remains at 41% however, so jump variable seems to impact only the shortest term horizon.

For SV regressions the adjusted R^2 is, as in the case for BPV, also quite high already for the third scale at 93%. However, unlike for the BPV case, including the jump variable in the regression does not improve the adjusted R^2

Table 5.9: Bipower Variance w/ Jumps Scales Regression - Portfolio(All)

	Dependent variable: Realized Vol.										
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	S10
rbpv1	−0.426***	0.953***	1.347***	1.540***	1.614***	1.727***	2.498***	1.646***	2.003***	1.977***	1.686***
	(0.030)	(0.024)	(0.012)	(0.009)	(0.005)	(0.010)	(0.014)	(0.015)	(0.029)	(0.013)	(0.016)
rbpv5	−1.315***	−0.453***	−1.132***	−0.662***	−0.658***	−0.770***	−1.700***	−0.644***	−1.073***	−1.032***	−0.676***
	(0.171)	(0.100)	(0.019)	(0.009)	(0.005)	(0.012)	(0.017)	(0.018)	(0.035)	(0.016)	(0.020)
rbpv22	−5.374***	−8.528***	−0.021	−0.152***	0.016***	0.022***	0.211***	−0.0001	0.073***	0.059***	−0.006*
	(0.718)	(0.415)	(0.088)	(0.017)	(0.002)	(0.003)	(0.003)	(0.004)	(0.007)	(0.003)	(0.004)
rjmp1	−0.782***	0.018	0.732***	0.901***	0.995***	1.008***	0.999***	0.999***	0.999***	0.998***	0.998***
	(0.021)	(0.028)	(0.015)	(0.010)	(0.003)	(0.002)	(0.001)	(0.001)	(0.0005)	(0.0001)	(0.0001)
Constant	−0.000	−0.000	0.000	−0.000	−0.000	−0.000	0.000	0.000	0.000***	0.000***	0.00000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.576	0.414	0.922	0.983	0.999	1.000	1.000	1.000	1.000	1.000	1.000
Note:											*p<0.1; **p<0.05; ***p<0.01

for the first scale which remains at 62%. It does however only slightly improve the adjusted R^2 for the second scale from 47% to 52%.

This result is similar to that obtained by Andersen *et al.* (2007) that found that jumps did not add much predictive value on future volatility and Patton & Sheppard (2015) showed that this was due to not including signs on the jumps. A possible is that by not including signs we're losing information that would be relevant for predicting future volatility.

5.5 Forecasting

In terms of forecasting we analyze and compare the forecast performance using the various proxies for volatility which have already been estimated in previous sections, as well as comparing between estimating using a normal time series vs. using wavelet scale decomposed time series. This will provide an indication whether there is a difference in the forecast performance between using the different volatility proxies as well as determining if there is an improvement in using wavelet decomposition. The result from this section should help in answering the first hypothesis whether wavelet decomposition helps to improve the predictive ability of the various volatility measures.

In the paper all the regular measures for forecast performance are provided however we focus on the MAPE and predictive r-squared for the analysis as

Table 5.10: Semi-Variances w/ Jumps Scales Regression - Portfolio(All)

	Dependent variable:										
	Realized Vol.										
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	S10
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
rpsv1	-0.367^{***}	0.842^{***}	1.162^{***}	1.333^{***}	1.470^{***}	1.555^{***}	1.623^{***}	1.551^{***}	1.625^{***}	1.648^{***}	1.648^{***}
	(0.034)	(0.034)	(0.015)	(0.009)	(0.008)	(0.006)	(0.005)	(0.005)	(0.003)	(0.004)	(0.003)
rpsv5	-1.696^{***}	-0.384^{***}	-0.611^{***}	-0.498^{***}	-0.531^{***}	-0.579^{***}	-0.653^{***}	-0.564^{***}	-0.655^{***}	-0.683^{***}	-0.682^{***}
	(0.165)	(0.087)	(0.020)	(0.009)	(0.009)	(0.008)	(0.006)	(0.006)	(0.004)	(0.005)	(0.004)
rpsv22	-7.889^{***}	-7.860^{***}	-1.236^{***}	-0.201^{***}	0.032^{***}	0.015^{***}	0.030^{***}	0.013^{***}	0.029^{***}	0.035^{***}	0.034^{***}
	(0.731)	(0.388)	(0.115)	(0.023)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
rnsv1	-0.547^{***}	0.840^{***}	1.274^{***}	1.324^{***}	1.588^{***}	1.667^{***}	1.657^{***}	1.737^{***}	1.657^{***}	1.637^{***}	1.640^{***}
	(0.037)	(0.029)	(0.015)	(0.010)	(0.009)	(0.007)	(0.005)	(0.005)	(0.003)	(0.004)	(0.003)
rnsv5	-0.931^{***}	-0.780^{***}	-0.702^{***}	-0.458^{***}	-0.650^{***}	-0.717^{***}	-0.697^{***}	-0.793^{***}	-0.694^{***}	-0.669^{***}	-0.675^{***}
	(0.205)	(0.117)	(0.026)	(0.010)	(0.010)	(0.008)	(0.007)	(0.006)	(0.004)	(0.005)	(0.004)
rnsv22	-2.477^{***}	-6.843^{***}	0.314^{**}	-0.217^{***}	0.060^{***}	0.054^{***}	0.040^{***}	0.057^{***}	0.037^{***}	0.033^{***}	0.034^{***}
	(0.868)	(0.528)	(0.135)	(0.025)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
rjmp1	-0.062^{**}	-0.565^{***}	-0.329^{***}	-0.047^{***}	0.008^{***}	0.005^{***}	-0.002^{***}	-0.001^{***}	0.0001^{***}	-0.00002^{***}	-0.0001^{***}
	(0.031)	(0.039)	(0.023)	(0.010)	(0.003)	(0.001)	(0.0002)	(0.0001)	(0.00002)	(0.00001)	(0.00000)
Constant	-0.000	-0.000	0.000	0.000	0.000	-0.000	0.000	0.000	-0.000	-0.000^{*}	0.000^{***}
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.621	0.516	0.935	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000
Note:											*p<0.1; **p<0.05; ***p<0.01

these are much more easily comparable between the different time series and samples; and additionally are not influenced by the magnitudes of the values in the time series. Using these measures are not without their disadvantages and one of the issues arising from using the MAPE is discussed in more detail in the paper, however as this paper is concerned with evaluating various models the ease of comparison is a major factor in favor of using these measures. Firstly a comparison is then made between in-sample and out of sample forecast performance for each model. Secondly the comparison between regular time series and the scale-decomposed series and finally a comparison between models is made. For clarity it's worth pointing out that with scale-decomposition we're less interested in the scale-specific data and more interested in the "horizon" data at the given scale (this includes both the details at the given scale plus the smooth).

In terms of 1-year ahead forecasting of realized volatility with the non-wavelet transformed time series (overall portfolio), there is very little difference in performance between the four methods. As shown in table 5.11 the MAPE varies from a low of 1.00 for the bipower variance estimator to a high of 1.09 for the semi-variance volatility estimator. So in regards to their forecast power

Table 5.11: Year Ahead Forecast Performance (Base Series)

Type	Statistic	BPV	BPV w/ Jumps	Semi-Variance	SV w/Jmp	BPV w/S.Jmp
Forecast	MAE	7.99e-05	7.96e-05	9.50e-05	9.00e-05	8.12e-05
	MAPE	1.00e+00	1.04e+00	1.02e+00	1.09e+00	1.00e+00
	RMSE	4.61e-04	4.59e-04	4.89e-04	4.73e-04	4.64e-04
	pred.rsq	4.57e-01	4.55e-01	4.64e-01	4.39e-01	4.48e-01
Regression	MAE	8.40e-05	8.47e-05	8.88e-05	8.46e-05	8.41e-05
	MAPE	8.81e-01	9.09e-01	8.21e-01	8.32e-01	8.78e-01
	RMSE	2.82e-04	2.80e-04	2.92e-04	2.81e-04	2.81e-04
	r.sq	5.47e-01	5.52e-01	5.13e-01	5.48e-01	5.48e-01
	adj.rsq	5.47e-01	5.51e-01	5.11e-01	5.47e-01	5.47e-01

then all the estimators are performing on equal basis and no one model really stands out. Compared to the MAPE from the regression (ranging 0.88 to 0.91) the difference is only about 10% degradation in performance. In absolute terms however the MAPE results are not really very good as the size of the error is as big or larger than the actual value. And in terms of predictive r-squared the values are also fairly close together ranging from 0.44 to 0.57 so performance quite evenly matched as well. These figures compare quite well to the adjusted r-squared from the regression which range from 0.51 to 0.55 the drop is quite limited. So overall the predictive power of the models do not degrade too much when switching from in-sample to out-of-sample judging by the 5% and 10% drops in r-squared and MAPE respectively, however the absolute level of the error is not desirable.

Also noteworthy from the results is that in terms of the MAPE statistic inclusion of the jump variable actually hurt forecasting performance marginally. For both bipower and semi-variance estimator the MAPE increased by around 4% and 7% respectively. So adding the jump variable is contributing to overfitting the model even though the adjusted R^2 showed a marginal improvement in the regression. Not much should be read into this result however since the other three statistics all show a marginal improvement when the jump variable is included. One of the reasons that the MAPE could sometimes show a bias is due to the fact that when the actual value is really small then the percentage error can still be quite large even when the size of the error is not extraordinary in any way.

With the scales forecasts we restrict the comparison to using only the MAPE which also allow following up on the question whether addition of the jump vari-

Table 5.12: Year Ahead Forecast MAPE Performance (w/ Scales)

	Forecast					Regression				
	BPV	**BPV w/Jmp**	**SV**	**SV w/Jmp**	**BPV w/ S.Jmp**	**BPV**	**BPV w/Jmp**	**SV**	**SV w/Jmp**	**BPV w/ S.Jmp**
D1	2.50712	2.20e+00	2.83e+00	2.83e+00	2.50061	20.51719	2.47e+01	2.14e+01	2.16e+01	13.75940
D2	4.01746	4.05e+00	5.10e+00	4.37e+00	3.99715	2.57022	2.57e+00	4.37e+00	4.18e+00	2.55961
D3	2.62670	2.47e+00	2.73e+00	2.54e+00	2.59798	2.68538	2.52e+00	2.08e+00	2.03e+00	2.65794
D4	0.59493	2.59e-01	2.27e-01	2.17e-01	0.60853	1.61118	9.59e-01	5.00e-01	4.86e-01	1.65124
D5	0.62085	2.39e-01	1.20e-01	1.23e-01	0.74209	1.30991	3.35e-01	1.22e-01	1.20e-01	1.36337
D6	0.36489	7.34e-02	2.00e-02	2.00e-02	0.67943	1.49861	1.67e-01	2.06e-02	2.03e-02	1.59764
D7	0.33671	6.20e-02	3.18e-03	4.88e-03	0.33297	1.95078	1.65e-01	1.69e-02	1.62e-02	1.92773
D8	2.70847	2.13e-01	6.41e-03	2.71e-03	1.02098	0.55421	2.00e-02	7.56e-04	8.45e-04	0.43245
D9	0.17475	2.02e-03	2.38e-05	2.28e-05	0.06207	0.18196	6.34e-03	8.98e-05	8.99e-05	0.13799
D10	0.12500	4.32e-03	2.56e-05	1.93e-05	0.95315	0.75654	7.72e-04	2.30e-05	2.18e-05	0.37740
S10	0.00273	2.82e-05	6.54e-07	5.45e-07	0.00418	0.00136	5.64e-06	1.67e-07	1.41e-07	0.00139

able is in fact contributing to over-fitting of the model and leading to a worse performance of same models without the jump variable.

Table 5.12 shows the 1-year ahead forecast performance as measured by the MAPE statistic for the different models and across different scales as well as comparing between regression and forecasting performance. Starting from the fourth scale upwards the MAPE is well below the 100% level, a significant improvement as compared to the non-wavelet transformed series. Moreover, the figures are declining rapidly as one goes up the scale which indicates a massive reduction in the noise and variability and the smoothing out of the signal over longer horizon timeframes.

Table 5.13: Year Ahead Cum. Forecast (w/ Scales)

Cum.Scale	BPV	BPV w/Jmp	SV	SV w/Jmp	BPV w/S.Jmp
1 - 10	0.6595	0.466900	5.54e-01	5.34e-01	0.9386
2 - 10	0.4557	0.281071	3.18e-01	3.02e-01	0.6986
3 - 10	0.3582	0.159272	1.66e-01	1.66e-01	0.5837
4 - 10	0.3527	0.070915	8.92e-02	8.97e-02	0.5789
5 - 10	1.1095	0.089521	3.84e-02	3.56e-02	1.8058
6 - 10	0.1729	0.010445	3.23e-03	3.39e-03	0.3804
7 - 10	0.1708	0.005716	6.31e-04	7.35e-04	0.2095
8 - 10	0.1036	0.002539	9.48e-05	8.44e-05	0.0310
9 - 10	0.0524	0.000797	8.28e-06	7.82e-06	0.0170
10	0.0107	0.000258	2.36e-06	1.91e-06	0.0463

The final part is in putting together the forecast series from the different scales into a combined forecast series. Table 5.13 shows the forecast performance statistics of scale-horizon data. The last row in the table shows the results for scale 10 detail and the rest of the low frequency components. The next row includes the results for next scale plus all the lower frequency components, all the way up to the 1st row in the table which includes the 1st scale detail plus all the low frequency components. Focusing once more on the MAPE, we see that when forecasting only on the highest scale horizon the error is only 1% or less (except for the BPV with signed jumps) and the percentage error gradually rises as more scales are included.

Starting with the bipower estimator, even with all scales included the percentage error on the BPV estimator is at 66% without the jump variable and with the jump variable included has only 47% error, which is a 20% improvement further validating inclusion of the jump variable in the estimation. Compared with the MAPE from forecasting of the original series which had 100% and 104% for the BPV and BPV with jumps respectively, this is a significant reduction in the forecast error and specifically in the case of BPV with jumps the error has more than halved in size as a result of employing wavelet decomposition.

A similar result is observed for the semi-variance estimators. The percentage error rises from a very low base as more scales are added to the longest horizon frequency. After all scales are added the error stands at 55% and 53% for the semi-variance and SV with jumps estimators respectively. Compared to the 95% and 90% for the SV and SV with jumps estimators on the original untransformed series, this represents an approximately 40% reduction in the error rate which is quite significant. As a side note we observe that the jump variable does not contribute a great deal to the forecast performance of the semi-variance estimator.

When looking to the BPV with signed jumps regression however, the improvement is only marginal. In the estimation and regression on the original time the MAPE were 0.88 and 1.00 respectively. With the wavelet transformed series, after including all the scales, the forecast error reduces to just 0.94 which is a much more modest gain as compared to the more than 50% reduction in forecast MAPE for the BPV with jumps.

Chapter 6

Empirical Results - Cross-Section

In the cross-section chapter investigate the question whether aggregate market volatility is a priced factor within the Fama-French three-factor model. We also investigate whether stocks with a high sensitivity to aggregate volatility would behave like a hedge against market volatility. And finally we investigate whether using wavelet decomposition might improve the accuracy of the analysis. As we have already found in this paper that risk as measured by the volatility level is different at different horizons, we do not discuss further the scale-specific betas for aggregate volatility and instead focus the discussion on the pricing of aggregate volatility risk. Here we find that aggregate volatility is indeed priced and is also behaving as a type of hedge against aggregate volatility. Additionally, we find that high sensitivity to idiosyncratic does not provide the hedge value that sensitivity to aggregate volatility does. And the results we discuss below.

6.1 Factor Pricing

To evaluate whether the moments can help in pricing risk in the Fama-French three-factor framework we follow a two-stage process. In the first stage we estimate the stock sensitivity to each of the factors and in the second we estimate the pricing of each of the factors. For this analysis we restrict our effort to the additional factors of aggregate market volatility as well as idiosyncratic volatility. These factors are chosen as they also tie into the question whether aggregate market volatility is priced and if some of that would show up in idiosyncratic volatility.

Results for the regression on the original non-wavelet transformed series

Table 6.1: Factor Pricing Comparison

	Dependent variable: Mean Returns				
	Factor Pricing Models in Fama-French 3-Factor Framework				
	Base	w/AggVol	w/Agg&IdioVol	w/AggVolHorizon	w/AggHoriz&IdioVol
	(1)	(2)	(3)	(4)	(5)
VolRisk			-0.001		0.0003
			(0.003)		(0.003)
AggVolInn0		0.563^{**}	0.522^{*}		
		(0.265)	(0.280)		
AggVolIShort0				-0.006	-0.006
				(0.011)	(0.011)
AggVolIMed0				0.087^{***}	0.087^{***}
				(0.027)	(0.028)
AggVolILong0				0.090	0.096
				(0.308)	(0.313)
MktExcRf0	-0.020^{***}	-0.018^{***}	-0.016^{***}	-0.019^{***}	-0.019^{***}
	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
SMB0	0.027^{***}	0.027^{***}	0.028^{***}	0.029^{***}	0.029^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
HML0	-0.020^{***}	-0.020^{***}	-0.019^{***}	-0.019^{***}	-0.019^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.001^{***}	0.001^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Observations	433	433	433	433	433
Adjusted R^2	0.299	0.304	0.303	0.314	0.313
Note:					*p<0.1; **p<0.05; ***p<0.01

are presented in table 6.1. The column (2) in the tables shows that the aggregate volatility factor is statistically (5% level) and economically significant in the pricing of factors in expected returns. The adjusted r-squared of 30% is also showing a relatively good fit and is approximately in line with the basic Fama-French factor pricing regression in column (1). In column (3) we show the pricing regression with aggregate volatility but this time controlling for idiosyncratic volatility and we see that aggregate volatility is still statistically and economically significant. In fact the beta coefficient is nearly unchanged from the basic regression without controlling for idiosyncratic volatility (2). This would suggest that aggregate volatility does not show up in idiosyncratic volatility. As an additional note which is not shown in this table is that in a pricing regression with only idiosyncratic volatility from the Fama-French time series regression, idiosyncratic volatility is not priced.

Table 6.2: Aggregate Vol. Factor Loading By Scale

	Dependent variable: Mean Returns										
	Factor Pricing Models										
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	S10
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
MktExcRf0	0.0002^{***}	0.0001^{**}	-0.0004^{***}	-0.001^{***}	0.0001	-0.0004^{***}	-0.0003^{**}	0.0004^{***}	0.0002^{***}	-0.0001^{***}	-0.010^{***}
	(0.00005)	(0.00004)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00003)	(0.002)
SMB0	0.0001	-0.00002	0.0001	0.00003	0.0001^{*}	0.0004^{***}	-0.001^{***}	0.001^{***}	-0.0002^{***}	-0.0001^{***}	0.003^{***}
	(0.00004)	(0.00003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00003)	(0.00001)	(0.001)
HML0	-0.0001^{***}	0.00004^{*}	0.0001^{***}	0.0002^{***}	0.00003	-0.0001	-0.0002^{***}	-0.0003^{***}	0.0002^{***}	0.0002^{***}	-0.002^{***}
	(0.00002)	(0.00002)	(0.00003)	(0.00005)	(0.00005)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00002)	(0.001)
AggVolInn0	-0.0002^{**}	0.0003^{*}	0.001^{***}	0.002^{***}	0.001	-0.001	-0.006	0.004	-0.012^{**}	0.019^{***}	1.306^{***}
	(0.0001)	(0.0001)	(0.0004)	(0.001)	(0.002)	(0.003)	(0.006)	(0.008)	(0.006)	(0.003)	(0.280)
Constant	0.00000^{***}	-0.00000	0.00000	0.00000^{***}	-0.00000	0.00000	-0.00001^{***}	-0.00000^{**}	0.00000^{***}	0.00000^{***}	0.0005^{***}
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00002)
Observations	433	433	433	433	433	433	433	433	433	433	433
Adjusted R^2	0.053	0.029	0.104	0.184	0.007	0.043	0.077	0.206	0.224	0.748	0.104

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

6.2 Scale-Decomposed Factor Pricing

Factor premiums for the aggregate volatility at different scales are shown in table E.1. The data shows that aggregate volatility is statistically significant for the lowest scales as well as for the highest scales. It is however not statistically significant for the middle scales (4 - 8) which correspond to the horizon 16 - 256 days (approximately 1 - 12 months).

A closer analysis of the coefficients for aggregate volatility reveals a similar pattern observed in the earlier regressions. The magnitude of the coefficients is continuously increasing as the scales are increasing, thus even the premium associated with volatility risk is more important for the higher scales than the lower scales. And the key insight therefore is that both the level of risk as well as the risk premium are more important at the higher scales than the lower scales.

The main anomaly in the data is that for the horizons in the middle where the premium is not statistically significant. A possible explanation could be that there may be some seasonalities in the data which are bringing down the correlations. An investigation of possible seasonalities however is beyond the scope of this paper. Alternatively this could simply be a result of a statistical anomaly whereby as we already saw in chapter 5 on volatility that volatility is less important for the lower scales but when it becomes important then it is much more important at the lowest scales than the other scales.

6.3 Risk-Sorted Returns and Forward Returns

The proposition is that if aggregate market volatility is a priced factor, then stocks with high sensitivity to market volatility could be considered as a type of hedge against market turmoil and would therefore be bid up in price resulting in lower subsequent returns. Our approach is in each period to sort the stocks into five portfolios according to their sensitivity to aggregate volatility. Then to calculate for each portfolio the contemporaneous average return as well as the next period average return. As a robustness check we also compare with different samples. And as a counter check to see that we didn't get a spurious result we also perform the exercise with portfolios sorted according to sensitivity to idiosyncratic volatility to see if we get a sensible result.

Table 6.3: Aggregate Volatility Sort

Portf..Risk	Avg..AggVolInn.Risk	Avg..Ret	Avg..Fwd.Ret
Lowest	-9.22e-05	0.000244	0.000983
L2	-3.95e-05	0.000389	0.000834
L3	-1.79e-05	0.000402	0.000841
L4	3.40e-06	0.000364	0.000637
Highest	3.82e-05	0.000376	0.000667

The average returns and forward returns, computed over the entire sample and sorted into portfolios based on sensitivity to innovations in aggregate volatility are given in table 6.3. Stocks that are least correlated with the aggregate volatility innovations are sorted in the lowest portfolio risk and those most positively correlated with aggregate volatility are sorted into the highest portfolio risk category. The stocks that had the highest sensitivity to innovations in aggregate volatility in the prior period generated lower returns in the next period than the stocks with the lowest sensitivity.

Table 6.4: Aggregate Volatility Medium Horizon Sort

Portf..Risk	Avg..AggVolIMed0.Risk	Avg..Ret	Avg..Fwd.Ret
Lowest	-6.51e-04	0.000286	0.001075
L2	-2.34e-04	0.000369	0.000586
L3	-4.26e-05	0.000335	0.000864
L4	1.61e-04	0.000423	0.000832
Highest	6.02e-04	0.000363	0.000605

As a robustness check we also evaluate the current period and forward returns sorting according to the medium horizon aggregate volatility which has been shown to be priced in the cross-section of returns. As for the case with the total aggregate volatility, the stocks with the highest sensitivity to medium horizon volatility had higher mean returns in the period followed by subsequent lower mean returns.

Table 6.5: Agg. Vol. Sort - Rolling Portfolio

	β_{aggVol}	Returns	Fwd.Returns
Lowest	-1.35e-03	0.000235	0.000738
L2	-4.09e-04	0.000366	0.000512
L3	1.13e-05	0.000422	0.000659
L4	4.33e-04	0.000427	0.000496
Highest	1.28e-03	0.000308	0.000598

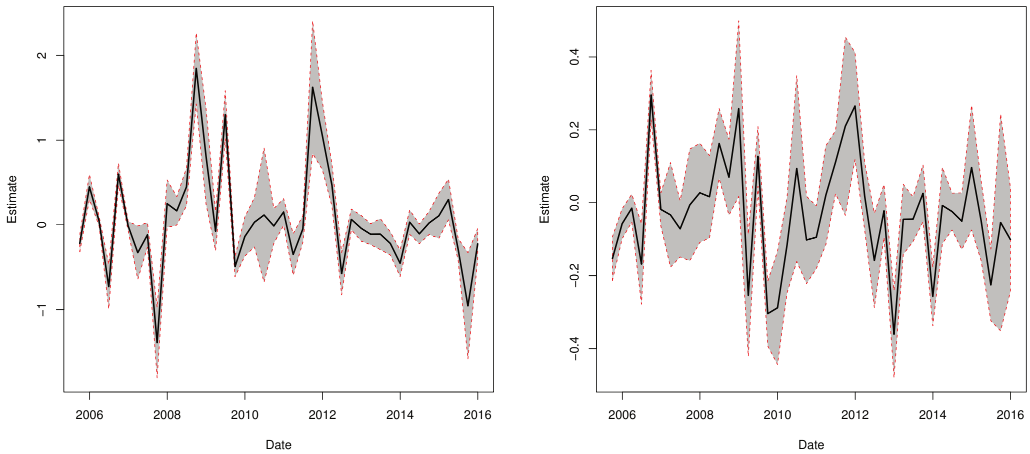
This data supports the proposition that the highest sensitivity portfolio was bid up during the period thus producing the higher mean returns, but it subsequently produced lower returns in the next period. This result supports the idea that stocks with a high sensitivity to aggregate volatility may behave like hedges so that they're bid up during times of higher volatility resulting in lower future returns.

Table 6.6: Idiosyncratic Volatility Risk Sort - Rolling Portfolio

	Volatility	Returns	Fwd.Returns
Lowest	7.77e-03	4.55e-04	4.07e-04
L2	1.02e-02	4.62e-04	6.58e-04
L3	1.23e-02	4.08e-04	5.60e-04
L4	1.52e-02	3.92e-04	6.55e-04
Highest	2.31e-02	4.07e-05	7.25e-04

The results obtained from a rolling sort based on idiosyncratic volatility are shown in table 6.6. As expected for stocks with the high sensitivity to idiosyncratic volatility the effect is not observed. In this case stocks with the highest sensitivity to idiosyncratic volatility in the prior period generate higher returns in the current period as well as in the next period than the stocks with the lowest sensitivity. This is a case of higher risk and higher reward. However, in terms of standard financial theory investors ought not to be compensated for taking on idiosyncratic risk. This result suggests therefore that there may be a part of idiosyncratic risk that is undiversifiable due to any number of market frictions for which investors demand compensation.

To test significance of the rolling regressions for the aggregate volatility premium the estimates together with the confidence bands are plotted in figure 6.1. In the figure are shown the estimated factor pricing for the total aggregate volatility as well as the medium horizon aggregate volatility.



(a) Aggregate Vol. Estimate (b) Medium Term Vol. Est.

Figure 6.1: Rolling Agg. Volatility Premia

From the diagrams it can be observed that total aggregate volatility estimate has a narrower confidence interval band and appears visually to be less variable than the medium horizon aggregate volatility estimate. Both estimates are hovering near zero with confidence bands overlapping the zero line so the factor premium is for the majority of the time not statistically significant. However, during times of market uncertainty (e.g. 2008) we see that the premium changes sharply and is strongly significant. So from this picture one could conclude that for the most part regular aggregate volatility is not generally priced however the extremes are highly significantly priced.

Chapter 7

Conclusion

This thesis set out to investigate whether the estimation and forecasting of the underlying volatility could be improved using realized volatility estimators within a HAR framework and whether wavelet transform could improve the horizon-specific results. In the cross-sectional analysis we sought to examine whether aggregate volatility was priced and furthermore, if assets with high sensitivity to aggregate volatility subsequently earned lower returns. That is, whether sensitivity to aggregate volatility might cause a security to become attractive as a hedge against market uncertainty. We employed multiresolution analysis to investigate the volatility structure and its evolution through time.

In terms of the volatility structure, looking at its energy content, we find that short term variability of volatility contributes much more to the overall variability of volatility with at least a fifth of the overall variability contributed by the 2-4 day time horizon. The effect is even more pronounced for the semi-variance estimators where the first scale contributes around a quarter of the variability. Furthermore the contribution is not constant through time and is higher during times of market uncertainty.

The multiresolution analysis helped improve quite significantly both the in- and out-of-sample estimation performance of all the estimators as compared to the original time series. We saw MAPE improvements of up to 50% with the BPV with jumps estimator when using the multiresolution data as compared to the original time series with the BPV estimator with signed jumps showing the least improvement.

All five realized volatility estimators had quite good performance generally and depending on the data series and the measurement statistic used sometimes there were mixed results. However, looking at just the forecasting and focusing purely on the mean absolute percentage error statistic, the BPV with jumps estimator performed the best among all the estimators and the BPV with signed jumps estimator was the worst performer overall. Inclusion of the unsigned jump variance to the semi-variances estimator had only a marginal improvement on the MAPE.

In the cross-section analysis we find that when evaluating over a long enough period aggregate volatility is priced. Over shorter periods the volatility premium is insignificant, particularly during quiet market periods. However, the premium rises very sharply during times of market uncertainty. So we could summarize that tail risk is priced but normal volatility is not.

Lastly, we also find that stocks that have a high sensitivity to aggregate volatility have lower future returns supporting the idea raised earlier that they become attractive contemporaneously as a hedge against market volatility.

As an area for future research would be to use as proxy for the aggregate volatility the realized market volatility instead of the implied volatility from the VIX in the cross-section of returns in order to use a more accurate measure of volatility.

Bibliography

Addison, P. S. (2002): *The illustrated wavelet transform handbook. Introductory theory and applications in science, engineering, medicine and finance.*

Andersen, T. G. & T. Bollerslev (1998): "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts*." *International Economic Review* **39(4)**: pp. 885–905.

Andersen, T. G., T. Bollerslev, F. X. Diebold, & H. Ebens (2001): "The distribution of realized stock return volatility." *Journal of Financial Economics* **61**: pp. 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, & P. Labys (2003): "Modeling and Forecasting Realized Volatility." *Econometrica* **71(2)**: pp. 579–625.

Andersen, T. G., T. Bollerslev, F. X. Diebold, P. Labys, & N. Mar (2007): *Modelling and Forecasting realized volatility*, volume 71.

Ang, A., R. J. Hodrick, Y. Xing, & X. Zhang (2006): "The Cross-Section of Volatility and Expected Returns." *The Journal of Finance* **LXI(1)**.

Barndorff-Nielsen, O. E., S. Kinnebrock, & N. Shephard (2009): "Measuring Downside Risk - Realized Semivariance." In "Volatility and Time Series Econometrics: Essays in Honor of Robert Engle," pp. 117–136. Oxford University Press.

Barndorff-Nielsen, O. E. & N. Shephard (2004a): "Econometrics of testing for jumps in financial economics using bipower variation." .

Barndorff-Nielsen, O. E. & N. Shephard (2004b): "Power and Bipower Variation with Stochastic Volatility and Jumps." *Journal of Financial Econometrics* **2(1)**: pp. 1—-37.

BARUNIK, J. & L. VACHA (2015): "Realized wavelet-based estimation of integrated variance and jumps in the presence of noise." *Quantitative Finance* **15(8)**: pp. 1347–1364.

BOLLERSLEV, T., S. Z. LI, & B. ZHAO (2016): "Good Volatility , Bad Volatility , and the Cross-Section of Stock Returns." *Working Paper* pp. 1–44.

CAMPBELL, J. Y. (1996): "Understanding Risk and Return." *The Journal of Political Economy* **104(2)**: pp. 298–345.

CHAMBERLAIN, G. & M. ROTHSCHILD (1983): "ARBITRAGE, FACTOR STRUCTURE, AND MEAN-VARIANCE ANALYSIS ON LARGE ASSET MARKETS." *Econometrica* **51(5)**: pp. 1281–1304.

CHAUDHURI, S. E. & A. W. LO (2016): "Spectral Portfolio Theory." *SSRN Electronic Journal* .

CHEN, J. (2002): "Intertemporal CAPM and the Cross-Section of Stock Returns."

CLEMENTS, A. E. & Y. LIAO (2013): "Modeling and forecasting realized volatility: getting the most out of the jump component."

COCHRANE, J. H. (2001): *Asset Pricing*. Princeton University Press.

CORSI, F. (2009): "A simple approximate long-memory model of realized volatility." *Journal of Financial Econometrics* **7(2)**: pp. 174–196.

DAUBECHIES, I. (1992): *Ten Lectures On Wavelets*, volume 61. Philadelphia: Society for Industrial and Applied Mathematics.

FAMA, E. F. & K. R. FRENCH (1992): "The Cross-Section of Expected Stock Returns." *The Journal of Finance* **47(2)**: pp. 427–465.

FAMA, E. F. & K. R. FRENCH (2016): "Dissecting Anomalies with a Five-Factor Model." *Review of Financial Studies* **29(1)**.

GENCAY, R., F. SELCK, & B. WHITCHER (2005): "Multiscale systematic risk." *Journal of International Money and Finance* **24**: pp. 55–70.

GENCAY, R., F. SELCUK, & B. WHITCHER (2001): "Scaling properties of foreign exchange volatility." *Physica A* **289**: pp. 249–266.

HARVEY, C. R. (2016): "... and the Cross-Section of Expected Returns." *Review of Financial Studies* **29(1)**: pp. 5–68.

LINDSAY, R. W., D. B. PERCIVAL, & D. A. ROTHROCK (1996): "The Discrete Wavelet Transform and the Scale Analysis of the Surface Properties of Sea Ice." *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING* **34(3)**.

MALLAT, S. G. (1989): "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11(7)**: pp. 674–693.

MERTON, R. C. (1973): "An Intertemporal Capital Asset Pricing Model." *Econometrica* **41(5)**: pp. 867–887.

PATTON, A. J. & K. SHEPPARD (2015): "Good Volatility, Bad Volatility: Signed Jumps and The Persistence of Volatility." *Review of Economics and Statistics* **97(3)**: pp. 683–697.

PERCIVAL, D. & A. T. WALDEN (2000): *Wavelet Methods for Time Series.*

PERCIVAL, D. B. & H. O. MOFJELD (1997): "Analysis of Subtidal Coastal Sea Level Fluctuations Using Wavelets." *Source Journal of the American Statistical Association* **92(439)**: pp. 868–880.

RAMSEY, J. B. (2002): "Wavelets in Economics and Finance: Past and Future." *Studies in Nonlinear Dynamics & Econometrics* **6(3)**.

ROSS, S. A. (1976): "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory* pp. 341–360.

SERROUKH, A. A., A. T. WALDEN, D. B. PERCIVAL, S. JOURNAL, A. STATISTICAL, N. MAR, A. SERROUKH, A. T. WALDEN, & D. B. PERCIVAL (2000): "Statistical Properties and Uses of the Wavelet Variance Estimator for the Scale Analysis of Time Series." *Journal of the American Statistical Association* **95(449)**: pp. 184–196.

TAMONI, A. (2011): "The multi-horizon dynamics of risk and returns." **1(March 2010)**.

Appendix A

Stock Ticker Information and Classification

Table A.1: Stock Description and Classification

Ticker	Name	Industry	MarketCap
BAC	Bank of America Corporation	Banks	$306.8B
C	Citigroup Inc.	Banks	$175.92B
JPM	J P Morgan Chase & Co	Banks	$372.49B
WFC	Wells Fargo & Company	Banks	$255.2B
AAPL	Apple Inc.	Technology	$823.61B
AMZN	Amazon.com, Inc.	Technology	$762.67B
CSCO	Cisco Systems, Inc.	Technology	$215.39B
IBM	International Business Machines Corporation	Technology	$134.46B
INTC	Intel Corporation	Technology	$246.69B
MSFT	Microsoft Corporation	Technology	$737.79B
ORCL	Oracle Corporation	Technology	$185.87B
CVX	Chevron Corporation	Oil	$241.97B
XOM	Exxon Mobil Corporation	Oil	$329.5B
KO	Coca-Cola Company (The)	Consumer	$184.7B
MCD	McDonald's Corporation	Consumer	$124.44B
PEP	Pepsico, Inc.	Consumer	$144.37B
DIS	Walt Disney Company (The)	Consumer	$149.21B
HD	Home Depot, Inc. (The)	Consumer	$214.94B
PG	Procter & Gamble Company (The)	Consumer	$183.09B
WMT	Walmart Inc.	Consumer	$257.72B
JNJ	Johnson & Johnson	Consumer	$344.1B
GE	General Electric Company	Industrials	$124.87B
SLB	Schlumberger N.V.	Industrials	$95.69B
CMCSA	Comcast Corporation	Communications	$147.74B
T	AT&T Inc.	Communications	$214.6B
VZ	Verizon Communications Inc.	Communications	$212.98B
QCOM	QUALCOMM Incorporated	Communications	$75.66B
MRK	Merck & Company, Inc.	Health	$160.14B
PFE	Pfizer, Inc.	Health	$220.1B

Appendix B

Variances Wavelet Decomposition

Table B.1: Variance Wavelet Decompostion - Full Sample

Type	Portfolio	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10	s10	waveVar	sigma.sq	Bias
	all	20.5	11.3	9.1	6.9	5.4	8.1	8.1	9.9	9.0	6.2	5.5	1.78e-07	1.78e-07	0
	banks	23.3	14.9	8.9	6.9	4.9	4.9	5.0	8.4	10.1	6.8	6.0	6.32e-06	6.32e-06	0
	tech	19.0	11.9	9.1	6.5	6.6	9.5	10.3	8.7	7.1	5.7	5.6	1.96e-07	1.96e-07	0
Realized Variance	oil	24.6	12.6	7.5	6.9	7.5	10.6	10.0	7.7	5.8	3.5	3.3	4.87e-07	4.87e-07	0
	cons	31.2	15.5	10.6	6.0	5.2	6.4	7.1	6.6	5.2	3.5	2.6	9.71e-08	9.71e-08	0
	indu	48.1	24.2	12.2	6.6	1.6	2.2	1.8	1.4	1.1	0.4	0.5	1.14e-05	1.14e-05	0
	coms	31.4	15.1	9.6	5.7	4.4	6.9	7.5	6.9	5.1	3.5	4.0	3.02e-07	3.02e-07	0
	all	19.5	10.6	7.3	7.0	6.5	8.5	8.9	10.5	9.2	6.4	5.5	8.78e-08	8.78e-08	0
	banks	12.0	10.0	9.0	8.6	6.8	5.1	4.8	10.8	13.6	10.5	8.7	1.70e-06	1.70e-06	0
	tech	13.6	8.1	5.7	6.7	7.9	10.9	11.8	11.3	9.5	7.3	7.1	8.79e-08	8.79e-08	0
BiPower Variance	oil	23.9	11.9	6.1	7.5	8.1	11.0	10.4	8.1	5.8	3.7	3.5	3.08e-07	3.08e-07	0
	cons	31.7	15.7	9.8	6.1	5.2	6.3	7.2	6.6	5.2	3.5	2.6	6.37e-08	6.37e-08	0
	indu	16.1	9.5	6.7	6.5	6.4	9.2	7.9	11.1	10.8	7.9	7.9	3.82e-07	3.82e-07	0
	coms	22.1	11.2	7.6	6.0	5.6	9.1	10.2	9.6	7.8	5.5	5.2	1.34e-07	1.34e-07	0
	all	39.4	19.4	11.0	5.9	3.3	4.2	3.7	4.1	4.0	2.6	2.5	3.39e-08	3.39e-08	0
	banks	41.9	23.6	10.1	4.3	2.7	3.3	3.3	3.4	3.7	2.0	1.8	2.85e-06	2.85e-06	0
	tech	42.1	21.6	12.7	6.1	3.9	3.7	3.6	2.1	1.6	1.3	1.3	5.00e-08	5.00e-08	0
Jump Variance	oil	40.3	18.7	11.6	6.7	5.0	4.9	4.2	3.1	2.7	1.5	1.3	4.66e-08	4.66e-08	0
	cons	44.0	18.5	11.4	4.9	4.1	4.0	3.7	3.5	2.7	1.8	1.5	6.80e-09	6.80e-09	0
	indu	50.0	25.0	12.5	6.6	1.3	1.8	1.4	0.8	0.4	0.1	0.1	1.08e-05	1.08e-05	0
	coms	48.4	23.6	12.2	6.4	2.9	1.7	1.8	1.3	0.6	0.4	0.7	1.02e-07	1.02e-07	0
	all	43.5	26.3	14.2	6.9	3.8	2.7	1.7	0.4	0.2	0.3	0.1	4.89e-08	4.89e-08	0
	banks	48.4	26.9	12.7	3.4	2.5	2.9	2.1	0.4	0.4	0.1	0.1	3.38e-06	3.38e-06	0
	tech	44.3	27.2	14.4	6.3	3.8	2.4	1.0	0.4	0.3	0.0	0.0	6.70e-08	6.70e-08	0
Signed Jump	oil	48.0	22.7	12.5	9.0	4.4	1.8	0.9	0.4	0.1	0.0	0.0	6.35e-08	6.35e-08	0
	cons	45.8	24.8	14.2	7.8	3.7	1.8	0.9	0.6	0.3	0.2	0.0	2.28e-08	2.28e-08	0
	indu	49.9	25.0	12.5	5.9	4.9	1.3	0.3	0.0	0.0	0.0	0.0	1.10e-05	1.10e-05	0
	coms	49.2	24.8	11.9	6.5	4.0	2.4	0.8	0.2	0.1	0.1	0.0	1.30e-07	1.30e-07	0
	all	24.1	14.3	10.0	6.4	4.6	7.4	7.7	7.6	7.6	5.5	4.9	5.67e-08	5.67e-08	0
	banks	36.8	19.9	10.0	4.5	3.5	4.1	3.8	4.4	5.8	3.7	3.4	3.34e-06	3.34e-06	0
	tech	24.1	15.0	8.9	5.2	5.7	8.6	8.7	7.4	6.4	5.0	5.0	5.71e-08	5.71e-08	0
Positive Semi-Variance	oil	28.1	11.4	6.7	5.8	7.2	11.1	10.2	7.4	5.5	3.4	3.2	1.29e-07	1.29e-07	0
	cons	27.5	14.3	10.5	5.4	5.1	7.3	8.5	7.4	6.0	4.4	3.7	1.86e-08	1.86e-08	0
	indu	48.8	24.5	12.3	6.3	3.2	1.8	1.1	0.8	0.7	0.2	0.3	4.79e-06	4.79e-06	0
	coms	35.1	17.0	9.9	6.4	4.8	5.6	5.8	5.1	3.9	3.0	3.4	9.03e-08	9.03e-08	0

Table B.2: Variance Wavelet Decompostion - Training Sample

Type	Portfolio	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10	s10	waveVar	sigma.sq	Bias
	all	18.1	9.8	8.3	6.6	5.4	8.7	8.9	11.0	10.0	7.3	6.1	1.74e-07	1.74e-07	0
	banks	23.3	14.8	8.9	6.9	4.9	4.9	5.1	8.5	10.2	7.0	5.5	6.89e-06	6.89e-06	0
	tech	16.5	10.4	8.2	6.1	6.8	10.2	11.3	9.7	7.9	6.8	6.2	1.91e-07	1.91e-07	0
Realized Variance	oil	24.0	12.1	7.1	6.7	7.6	10.9	10.3	8.0	5.9	3.9	3.6	5.11e-07	5.11e-07	0
	cons	29.1	13.7	9.5	5.4	5.2	7.2	8.3	7.9	6.1	4.3	3.4	8.74e-08	8.74e-08	0
	indu	48.2	24.2	12.2	6.6	1.6	2.2	1.8	1.4	1.1	0.4	0.4	1.26e-05	1.26e-05	0
	coms	30.0	14.3	9.2	5.5	4.4	7.3	8.2	7.5	5.5	4.1	4.1	3.02e-07	3.02e-07	0
	all	17.7	9.7	6.8	6.8	6.6	9.0	9.5	11.3	9.8	7.1	5.8	8.98e-08	8.98e-08	0
	banks	11.7	9.9	9.0	8.7	6.9	5.2	4.8	10.9	13.8	10.8	8.2	1.85e-06	1.85e-06	0
	tech	13.2	7.8	5.4	6.6	7.9	11.1	12.1	11.6	9.7	7.8	6.9	9.41e-08	9.41e-08	0
BiPower Variance	oil	23.8	11.8	6.0	7.5	8.2	11.1	10.4	8.2	5.8	3.8	3.5	3.35e-07	3.35e-07	0
	cons	27.8	13.4	8.7	5.6	5.5	7.4	8.9	8.3	6.5	4.5	3.5	5.50e-08	5.50e-08	0
	indu	13.6	8.3	6.1	6.5	6.6	9.8	8.5	12.1	11.7	8.8	8.1	3.88e-07	3.88e-07	0
	coms	19.2	9.8	6.9	5.8	5.8	9.9	11.3	10.7	8.7	6.3	5.6	1.32e-07	1.32e-07	0
	all	39.9	18.8	10.3	5.3	2.8	4.2	3.8	4.4	4.3	3.2	2.9	3.29e-08	3.29e-08	0
	banks	42.0	23.6	10.0	4.3	2.7	3.3	3.3	3.4	3.7	2.0	1.7	3.14e-06	3.14e-06	0
	tech	42.5	20.7	11.9	5.3	3.6	4.0	4.1	2.4	1.8	1.8	1.8	4.26e-08	4.26e-08	0
Jump Variance	oil	41.1	17.0	10.6	5.9	4.7	5.3	4.9	3.7	3.0	1.9	2.0	4.05e-08	4.05e-08	0
	cons	44.6	18.1	11.3	4.7	3.8	3.9	3.7	3.5	2.7	2.0	1.7	7.20e-09	7.20e-09	0
	indu	50.0	25.0	12.5	6.6	1.3	1.8	1.4	0.8	0.4	0.1	0.1	1.19e-05	1.19e-05	0
	coms	48.6	23.6	12.1	6.3	2.8	1.7	1.8	1.3	0.6	0.5	0.6	1.10e-07	1.10e-07	0
	all	44.9	26.2	13.8	6.5	3.3	2.7	1.8	0.2	0.2	0.2	0.1	4.00e-08	4.00e-08	0
	banks	48.6	27.0	12.6	3.3	2.5	2.9	2.1	0.4	0.3	0.1	0.1	3.68e-06	3.68e-06	0
	tech	45.9	27.3	14.0	5.9	3.3	2.2	0.9	0.3	0.2	0.0	0.0	5.82e-08	5.82e-08	0
Signed Jump	oil	50.2	21.4	12.0	9.2	4.1	1.7	1.0	0.3	0.0	0.0	0.0	5.51e-08	5.51e-08	0
	cons	49.9	24.2	13.4	8.0	2.6	1.1	0.3	0.3	0.1	0.0	0.0	1.27e-08	1.27e-08	0
	indu	49.9	25.0	12.5	5.9	4.9	1.3	0.3	0.0	0.0	0.0	0.0	1.21e-05	1.21e-05	0
	coms	49.6	24.9	11.6	6.5	4.0	2.3	0.7	0.2	0.1	0.0	0.0	1.26e-07	1.26e-07	0
	all	24.1	14.3	10.0	6.3	4.6	7.4	7.7	7.7	7.7	5.7	4.6	6.21e-08	6.21e-08	0
	banks	36.9	20.0	10.0	4.5	3.5	4.1	3.8	4.4	5.8	3.8	3.0	3.68e-06	3.68e-06	0
	tech	23.9	14.9	8.9	5.2	5.8	8.7	8.8	7.5	6.4	5.2	4.7	6.19e-08	6.19e-08	0
Positive Semi-Variance	oil	28.1	11.4	6.6	5.8	7.2	11.2	10.2	7.4	5.5	3.5	3.2	1.41e-07	1.41e-07	0
	cons	27.4	14.2	10.4	5.3	5.1	7.3	8.6	7.5	6.0	4.6	3.6	2.02e-08	2.02e-08	0
	indu	48.8	24.5	12.3	6.3	3.2	1.8	1.1	0.8	0.7	0.3	0.2	5.29e-06	5.29e-06	0
	coms	35.2	17.1	9.9	6.4	4.8	5.6	5.9	5.1	3.9	3.1	3.1	9.88e-08	9.88e-08	0

Table B.3: Variance Wavelet Decompostion - Testing Sample

Type	Portfolio	d1	d2	d3	d4	d5	d6	d7	s7	waveVar	sigma.sq	Bias
Realized Variance	all	39.2	22.9	15.4	9.3	6.0	3.9	2.2	1.1	2.17e-07	2.17e-07	0
	banks	37.0	23.4	16.7	10.1	6.0	3.4	2.2	1.1	6.26e-07	6.26e-07	0
	tech	37.6	23.3	16.1	9.7	5.8	4.0	2.4	1.1	2.48e-07	2.48e-07	0
	oil	35.5	22.4	15.0	10.2	6.3	5.2	3.4	2.0	2.68e-07	2.68e-07	0
	cons	40.8	23.4	15.4	8.9	5.6	3.4	1.8	0.8	1.90e-07	1.90e-07	0
	indu	42.0	23.9	14.0	8.3	5.2	3.4	2.0	1.0	4.58e-07	4.58e-07	0
	coms	45.8	23.0	13.8	7.7	4.7	2.8	1.5	0.7	2.98e-07	2.98e-07	0
BiPower Variance	all	42.6	22.4	14.4	8.9	5.3	3.6	1.9	0.9	6.75e-08	6.75e-08	0
	banks	45.0	22.0	14.1	8.5	4.8	2.9	1.7	0.8	1.94e-07	1.94e-07	0
	tech	30.1	19.9	17.3	12.2	8.0	7.1	3.7	1.7	2.60e-08	2.60e-08	0
	oil	32.2	17.8	14.2	11.3	6.5	7.3	6.5	4.3	4.57e-08	4.57e-08	0
	cons	45.8	23.9	13.9	7.8	4.3	2.6	1.3	0.6	1.47e-07	1.47e-07	0
	indu	45.7	23.5	13.5	7.6	4.4	2.8	1.6	0.9	3.15e-07	3.15e-07	0
	coms	46.8	23.1	13.6	7.7	4.2	2.7	1.3	0.6	1.50e-07	1.50e-07	0
Jump Variance	all	35.6	23.7	16.3	9.7	6.8	4.1	2.6	1.3	4.37e-08	4.37e-08	0
	banks	29.1	24.9	19.2	11.5	7.1	3.8	2.8	1.5	1.33e-07	1.33e-07	0
	tech	40.9	24.8	15.4	8.5	4.9	2.8	1.8	0.9	1.21e-07	1.21e-07	0
	oil	37.6	24.7	15.5	9.7	6.0	3.7	1.9	1.0	1.04e-07	1.04e-07	0
	cons	29.4	26.0	12.1	10.1	10.2	6.3	3.9	1.9	3.20e-09	3.20e-09	0
	indu	50.6	18.0	10.0	8.7	4.8	4.4	2.5	1.2	1.16e-08	1.16e-08	0
	coms	43.9	23.0	14.1	7.9	5.6	2.9	1.7	0.9	2.73e-08	2.73e-08	0
Signed Jump	all	39.9	26.6	15.3	7.9	5.4	2.8	1.4	0.7	1.32e-07	1.32e-07	0
	banks	37.9	26.2	16.5	9.2	5.4	2.5	1.5	0.8	4.70e-07	4.70e-07	0
	tech	38.5	27.0	16.0	7.9	5.7	3.3	1.2	0.4	1.50e-07	1.50e-07	0
	oil	40.2	27.6	14.6	8.3	5.4	2.4	0.9	0.5	1.42e-07	1.42e-07	0
	cons	41.8	25.6	15.1	7.7	5.0	2.5	1.5	0.8	1.18e-07	1.18e-07	0
	indu	42.4	28.4	14.5	7.0	4.3	2.1	1.0	0.5	2.84e-07	2.84e-07	0
	coms	46.6	24.5	13.5	6.7	4.4	2.5	1.3	0.5	1.69e-07	1.69e-07	0
Positive Semi-Variance	all	35.4	15.3	15.4	11.1	7.1	7.8	5.0	2.8	4.30e-09	4.30e-09	0
	banks	32.3	18.1	16.5	11.4	6.3	7.4	5.3	2.8	8.70e-09	8.70e-09	0
	tech	43.9	19.0	15.0	9.9	3.3	3.0	3.7	2.4	1.03e-08	1.03e-08	0
	oil	32.8	15.6	15.4	10.1	5.3	7.1	8.2	5.5	1.54e-08	1.54e-08	0
	cons	36.2	16.4	15.7	12.3	7.5	7.1	3.3	1.5	3.00e-09	3.00e-09	0
	indu	34.4	15.1	15.3	10.0	6.4	7.5	6.8	4.4	1.31e-08	1.31e-08	0
	coms	43.2	20.0	15.0	10.2	4.8	2.8	2.4	1.7	7.80e-09	7.80e-09	0

Appendix C

Additional Regressions

Table C.1: Bipower Vol. w/ Signed Jumps - Portfolios

	Dependent variable:						
	Realized Vol.						
	all	banks	tech	oil	cons	indu	coms
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
rbpv1	0.359***	0.766***	0.529***	0.015	0.143***	0.416*	0.134***
	(0.034)	(0.051)	(0.038)	(0.032)	(0.034)	(0.223)	(0.044)
rbpv5	0.704***	0.557***	0.618***	0.923***	0.556***	0.587	0.769***
	(0.058)	(0.080)	(0.059)	(0.055)	(0.061)	(0.361)	(0.076)
rbpv22	0.203***	0.268***	0.080*	0.148***	0.382***	0.189	0.233***
	(0.050)	(0.072)	(0.048)	(0.049)	(0.060)	(0.305)	(0.068)
rsjmp1	−0.061**	−0.050**	−0.085***	−0.183***	0.198***	−0.003	0.062**
	(0.029)	(0.020)	(0.024)	(0.048)	(0.049)	(0.021)	(0.026)
Constant	0.00003***	0.00003	0.00003***	0.00005***	0.00002***	0.0002*	0.00004***
	(0.00001)	(0.00004)	(0.00001)	(0.00001)	(0.00001)	(0.0001)	(0.00001)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.547	0.491	0.580	0.443	0.334	0.030	0.351
Note:							*p<0.1; **p<0.05; ***p<0.01

Table C.2: Bipower Variance w/ Signed Jumps Scales - Portfolio(All)

	Dependent variable:										
	Realized Vol.										
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	S10
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
rbpv1	−0.451***	0.951***	1.577***	1.688***	1.220***	2.120***	4.068***	8.396***	28.963***	−41.385***	50.530***
	(0.037)	(0.025)	(0.014)	(0.020)	(0.029)	(0.138)	(0.378)	(0.459)	(1.092)	(3.284)	(3.024)
rbpv5	−1.119***	−0.459***	−0.885***	−0.470***	−0.045	−0.622***	−2.953***	−9.044***	−34.250***	53.405***	−62.526***
	(0.208)	(0.101)	(0.026)	(0.022)	(0.031)	(0.167)	(0.469)	(0.566)	(1.348)	(4.054)	(3.733)
rbpv22	−9.702***	−8.408***	−0.565***	0.389***	−0.046***	−0.264***	0.201**	2.050***	6.751***	−10.616***	13.444***
	(0.858)	(0.418)	(0.120)	(0.036)	(0.016)	(0.041)	(0.097)	(0.109)	(0.258)	(0.771)	(0.710)
rsjmp1	−0.290***	−0.011	0.206***	0.068***	0.095***	0.248***	0.029*	−0.339***	−0.297***	−0.338***	−0.232***
	(0.022)	(0.021)	(0.014)	(0.013)	(0.012)	(0.014)	(0.017)	(0.013)	(0.015)	(0.008)	(0.021)
Constant	−0.000	−0.000	0.000	−0.000	−0.000	0.000	0.00000	−0.00000*	−0.00000**	0.00000***	0.00001***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Observations	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341	2,341
Adjusted R^2	0.376	0.414	0.858	0.919	0.947	0.986	0.984	0.998	0.999	0.999	1.000

Note: *p<0.1; **p<0.05; ***p<0.01

Appendix D

Horizon-Relevant Volatilities

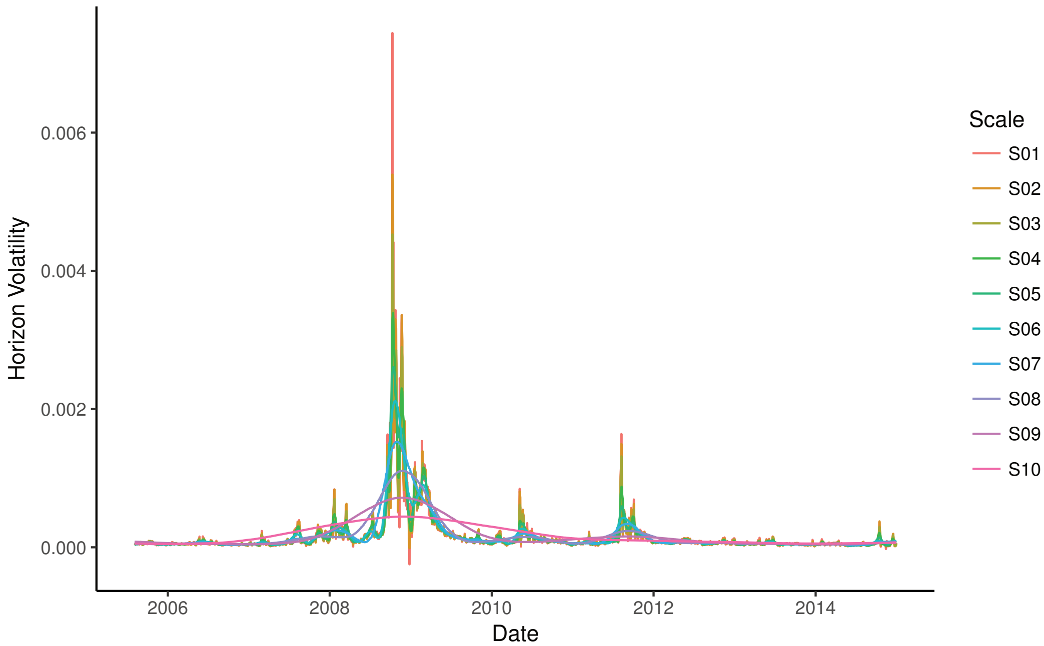


Figure D.1: BPV Horizon Volatility (All Scales)

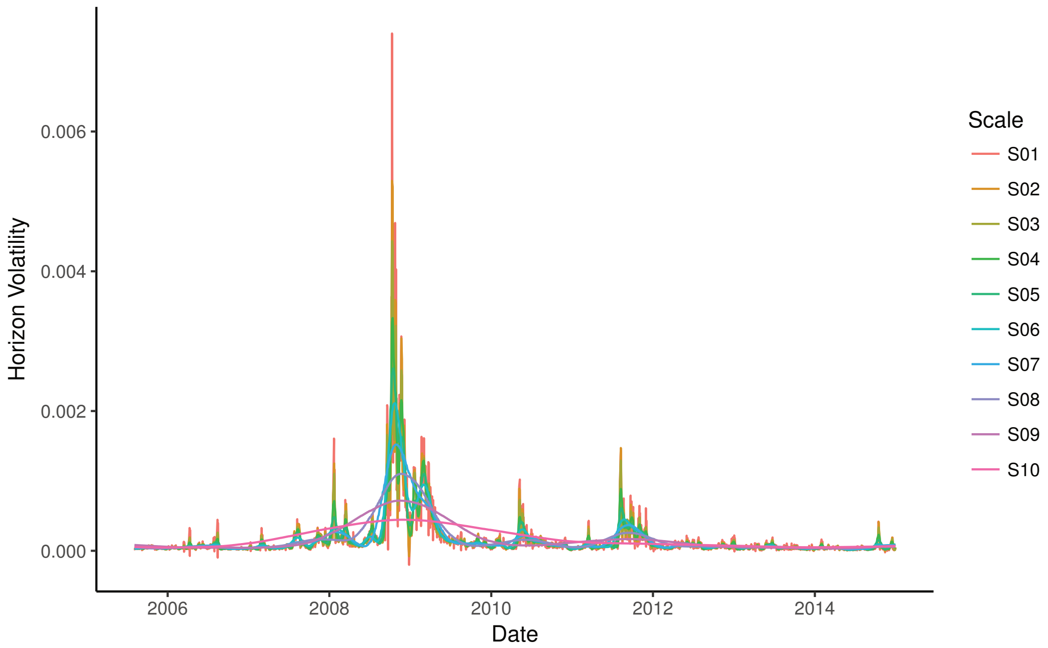


Figure D.2: BPV w/ Jumps Horizon Volatility (All Scales)

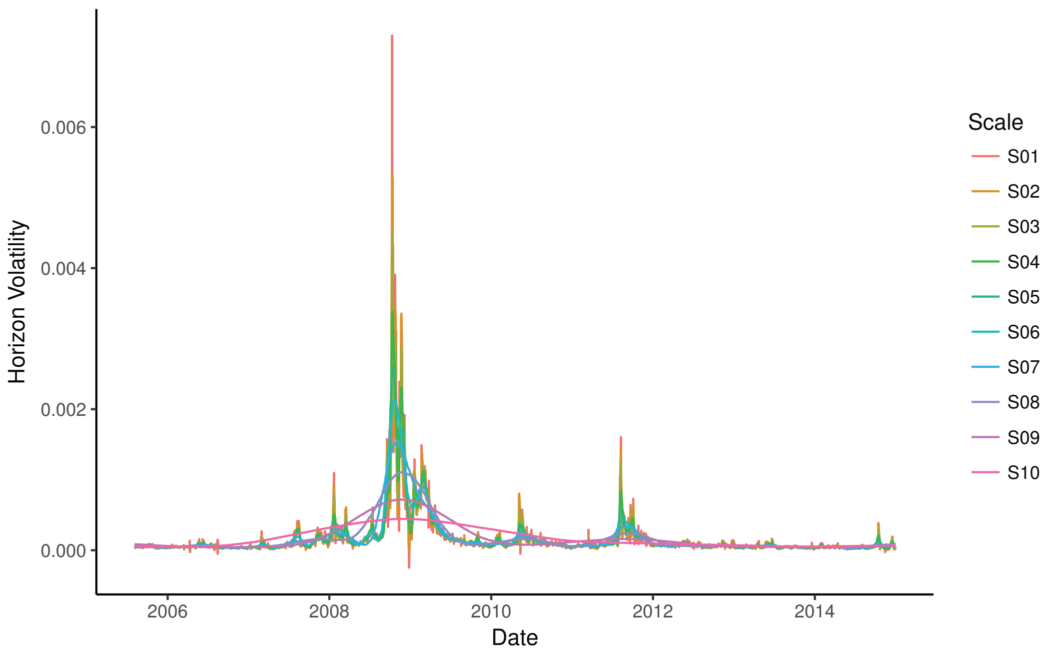


Figure D.3: BPV w/ Signed Jumps Horizon Volatility (All Scales)

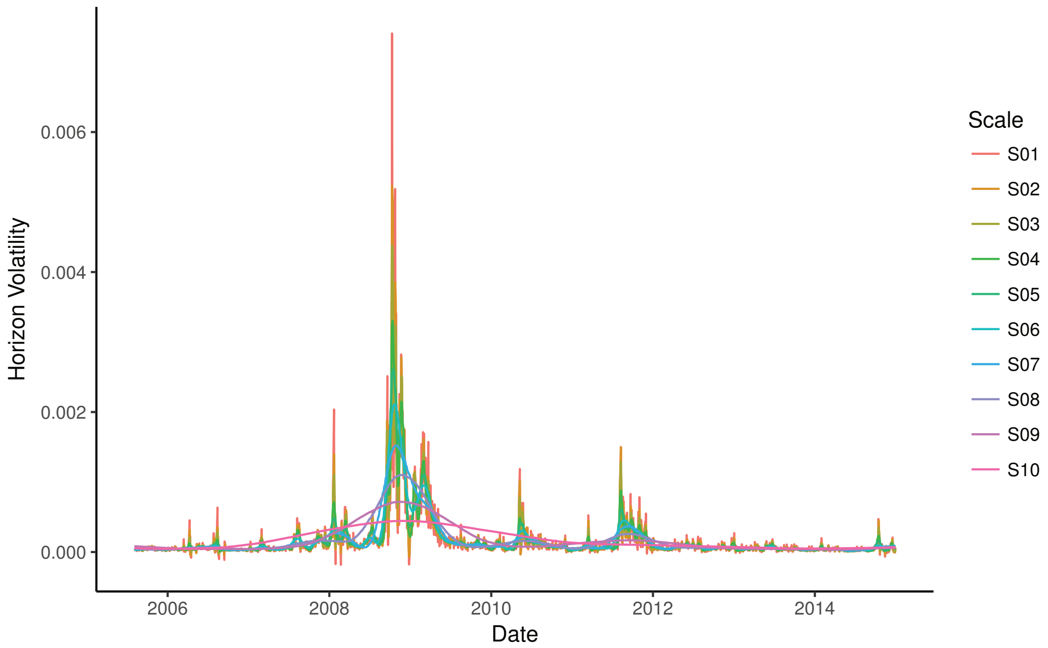


Figure D.4: SV Horizon Volatility (All Scales)

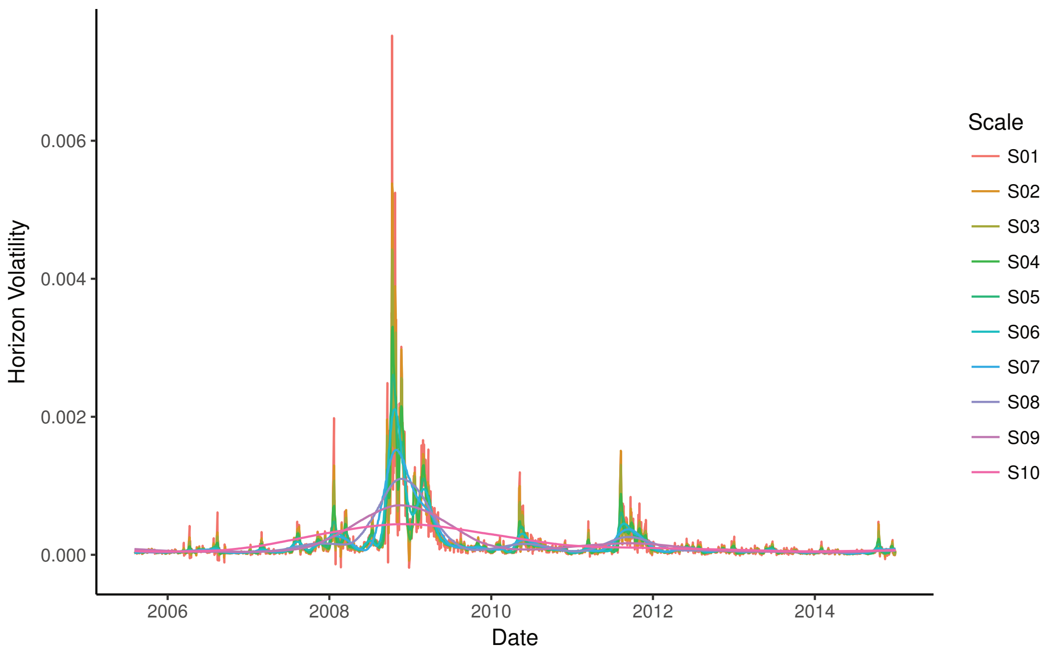


Figure D.5: SV w/ Jumps Horizon Volatility (All Scales)

Appendix E

Factor Loadings

Table E.1: Aggregate Vol. Factor Loading By Scale

	Dependent variable: Mean Returns										
	Factor Pricing Models										
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	S10
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
MktExcRf0	0.0002^{***}	0.0001^{**}	-0.0004^{***}	-0.001^{***}	0.0001	-0.0004^{***}	-0.0003^{**}	0.0004^{***}	0.0002^{***}	-0.0001^{***}	-0.010^{***}
	(0.00005)	(0.00004)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00003)	(0.002)
SMB0	0.0001	-0.00002	0.0001	0.00003	0.0001^{*}	0.0004^{***}	-0.001^{***}	0.001^{***}	-0.0002^{***}	-0.0001^{***}	0.003^{***}
	(0.00004)	(0.00003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00003)	(0.00001)	(0.001)
HML0	-0.0001^{***}	0.00004^{*}	0.0001^{***}	0.0002^{***}	0.00003	-0.0001	-0.0002^{***}	-0.0003^{***}	0.0002^{***}	0.0002^{***}	-0.002^{***}
	(0.00002)	(0.00002)	(0.00003)	(0.00005)	(0.00005)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00002)	(0.001)
AggVolInn0	-0.0002^{**}	0.0003^{*}	0.001^{***}	0.002^{***}	0.001	-0.001	-0.006	0.004	-0.012^{**}	0.019^{***}	1.306^{***}
	(0.0001)	(0.0001)	(0.0004)	(0.001)	(0.002)	(0.003)	(0.006)	(0.008)	(0.006)	(0.003)	(0.280)
Constant	0.00000^{***}	-0.00000	0.00000	0.00000^{***}	-0.00000	0.00000	-0.00001^{***}	-0.00000^{**}	0.00000^{***}	0.00000^{***}	0.0005^{***}
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00002)
Observations	433	433	433	433	433	433	433	433	433	433	433
Adjusted R^2	0.053	0.029	0.104	0.184	0.007	0.043	0.077	0.206	0.224	0.748	0.104

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$